

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAACACAGAAAACCTGTTAGAAATGTTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTATGTTGCTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCCTTGTAAGTGAATCTGAGTTGTTTAGGACTTCTATTGTGGCAAACTT
CCAGAAAAACAACCCCTTTTGTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCACCAAATGCAGGCCAAAATCCATGGCAACACAAGTC
TTCGTGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCTTCTTT
GGTTTTTCTGACTTACATTGCTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCTCTATGACACTGCACCTTGCCCTATTAAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCACAGGATTGGGGAAGG
TTCACAGAAGTTGCTTATCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAACATGAAAGAAGCCATTTGATAGATTATCTTAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATTAAGTCRAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVINWTSAAFI~~FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV~~
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFA~~NHVS~~GAVLTFG
MGSLYMFVQTILSYQM~~QPKIHGKQVFWIRLLLVINCGVS~~ALSM~~LT~~CSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWMSM~~SFS~~FFGFFLT~~YIRD~~FQKISLRVEANLHGLTLYDTAPCPINN~~ER~~TR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCGAGACC
GTGTGAGGGGGCCTGTGGCCCGAGCGTGTCTGGCCTCGGGAGTGGGAAGTGGAGGCGAGGCGCTTC
CTTACACTTGCCTGAGTTTCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTTTG
GATTGGGTGGCTTTCTTCATGGGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTCTTGCAACATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTCTACTGAAAAATGAACCTGTGTGAATCTGCTGATCTGCG
TTTTCATGGTGGCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACCAACGA
CTGCTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTCTGGAACCTAGGAGATCCCTTTCC
CATTTCTCAGCCCAAAACATGSGATCTTATCCATAGAACAGCTCATCAGCGGGTTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGTCAACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGSGAGTGTGCAAAACCATGGATATGAT
CATAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAAGGGGAAGTGATAACA
AACCATCAGGTTTCTGGGAATGATAAAAAGTGTACCACCTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAAGCAGGCGAGCTTTTCTGGAACAGCTGATCT
ATATGCTACCAAGGAGAGATAGAATACTCCAAAACCTTCAAGGGGAATATTTTAAATTTCTTGTT
ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTATGCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATATCTGGGAATCCAAATTTGATGT
GAAGTTTGGTCCCAACACATTTCTCTCATTTCTTGTGGAATATCATCTGCATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCCTCCAATGCTATGTCCTG
CTATTAGCACAGATAATGGGCAATGACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCCTT
AGAATACCGCACCATATCACTGAAGTCTTGAGAACTGCACTTCACTTCTATCACCGTTGTTTG
ATGIGATCTTCTGGTGCAGCGCTCTCTCATGATATCTTCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAATGGCACCTTGAACCTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTA
GATATAAGAGGGGGGAAAATGGAACCAAGGCGCTGACATTTTATAAACAAACAAATGCTATGGTAGC
ATTTTTCACCTCTATAGCATACTCCTTCCCGCTCAGGTGATACATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGATATGAGGCTGG
GTATAGAGCGGAGGAGGACCAAGAACTAAAGGTGAAAAATACTGGAACCTGGGGCAAGCATGT
CTATGGTAGCTGAGCCAAACAGTAGGATTTCCGTTTTRAGGTTACACATGGAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAATCAGAGACTGTAAACAAAAAAGGCGCGCCGCG
ACTCTAGAGTCGACCTGCAGAAAGCTTGGCCGCCATGGCCCACTGTTTATTGCGAGTTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGPGWLFMRQLFKDYEIRQYVQVIFSVTFAPSCTMFELIIFEILGV
LNSSSRYPFWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLFLSCLLWLTFFMYFFWKLGGP
FPILSPKHGILSIEQLISRVGIVGVTLMALLSGFGAVNCPITYMSYFLRNVTDDILALERLLQ
TMDMIISKRRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSNLTLIQQEVDAL EELSRO
LFLETADLYATKERIEYSKTFKGYFNFLGYFFSYICVWKIFMATINIVFDRVGKTDPTVRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYATISSKSSNVIVLLLAQIMGY
FVSSVLLTRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALESSILFLYLAHKQAPKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCCTGGTTATGAAGTGGAGCAGTGAAGTGTGAGCCTCAACATAGTTCC
 AGAAGCTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCCTGGCTCTGAAGGGGTAGGCACAGTGGCCAGTGTCTCAGCCTGGTGTGCTCTCAGCT
 TCCATCTGGACCAACGAGGCTCTCTGGTCAGGGCTCTTTGCGTGCAGAAAGCCTTTCCATCCAGGT
 GTCTATGCAGAAATTATGGGGATCACCTCTGTGAGCAAAAAGGCGAACAGCAGCTGAATTTACAG
 AAGCTAAGSAGGCTCTGAGGCTGCTGGGACTAAGTTTGGCGCGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTCTGGTCACTCTTAG
 GATTAGCCCAAAACCCAAAGTGTGGGAAAAATGGGGTGGTCTCTGATTGGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
 ATTATCACCAACAAAGATCCCATATTCAACACTCAAACTGCAACACAAACACAGAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACATACTGCCCCCTACTACTACTCCTCT
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTGTGTCAAGAAAGTTTATATG
 GAAACTAGCACCATGTCTACAGAACTGAACCAATTTGTTGAAATAAAGCAGCATTTCAAGATGA
 AGCTGTGGGTTTGGAGGTGCCCCACGGCTCTGCTAGTGTCTGCTCTCTCTCTTCTTGGTGTG
 CAGCTGGTCTTGGATTTTGCTATGTCAAAAGGTATGTGAAGGCCCTTCCCTTTACAAACAAGAT
 CAGCAGAAGGAAATGATCGAAACCAAGTAGTAAAGGAGGAGAGGCCAATGATAGCAACCCCTAA
 TGAGGAATCAAAGAAACTGATAAAACCCAGAAAGTCCAAAGAGTCCAAGCAAAACTACCGTGC
 GATGCTTGGAGCTGAAGTTTATGATGAGACAGAAATGAGGAGACACACTGAGGCTGGTTCTTT
 CATGCTCCTTACCTGCCCGAGCTGGGGAAATCAAAGGGCCAAAGAACCAAAGAGAAAGTCCA
 CCGTTGGTTCTTAAGTGAATCAGCTCAGGACTGCCTATGGACTATGGAGTGACCAAAAGAGAAAT
 GCCCTTCTCTTATTGTAACCTGTCTGGATCTATCCTCTACCTCCAAAGCTTCCACAGGCT
 TTCTAGCTCTGGCTATGTCTTAATATATCCCACTGGGAGAAAGGAGTTTGCAGAGTGCAGGAGC
 CTAAACATCTCATCAGTATCCAGTGGTAAAGGGCTCTGCTGTCTGAGGCTAGGTTGGGTTG
 AAAGCCAAAGGAGTCACTGAGACCAAGGCTTTCTTACTGATTCCGAGCTCAGACCCCTTTCTCA
 GCTCTGAAGAGAAACAGCTATCCCACTGACATGTCTTCTGAGCCCGTAAAGACAAAGAAAT
 GGCAGAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATACTTGAGACCTAATCTCTGTAAA
 GCTAAAAATAAGAAATAGACAAAGGCTGAGGATACGACAGTACACTGTACAGCAGGACTGTAAC
 ACAGACAGGCTCAAGTGTCTCTGACACACTGAGTTGGAATCACTGTTAGACACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATTTTCTCTAGGAAATATACCTTTACAAGTAA
 AAAATAAAACTCTTATATAATTCTATTTTATCTGAGTTACAGAAATGATTACTAAGSAGAT
 ACTCAGTAATTGTTAAAAAGTAAATAAATTCACAAACATTTGCTGAATAGCTACTATATGCT
 AAGTGTGTGCAAGGTATACACTCTGTAATTGAAATATTCTCTCAAAAAATGACACATAGTAG
 AACGCTATCTGGGAGCTATTTTTCTAGTTTGAATTTCTAGCTTATCTACTTTCCAAACTAAT
 TTTTATTTTTGTGAGACTAATCTTATTCATTTTTCTCTAATATGGCAACATTATAACCTTAAT
 TATTATTACATACCTAAGAGTACATTTGTACCTCTATATACCAAGCACATTTTAAAGTGGC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAAAGAAAGGAGCTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAGCATTTAGAAAACTT

FIGURE 6

MARCFSLVLLLSIWTTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTAEACRLLG
LSLAGKDQVETALKASFETCSYGVWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWNTSCIEPIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFEMETSTMSTETEPFVENKAAFKNEAAGFGVPTALLVLALLFFGAAGLGFCYVK
RYVKAFPTFNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCGCGGCTCCGGACCGCGCGCGCGCCACCGCGCGCTCCGCACTCGACCCGCGAGCCCGG
 GGCCGCCGCGGGGAGCGAGCAGATCCAGTCCGGCCCGCGAGCGCAACTCGGTCCAGTCCGGGGGG
 CGGCTCGGGCGCGAGAGCGGAGATGCGAGCGGTTGGGGCCACCTTGCTGTGCTGTGTGCGGG
 CGCGCGTCCCGACGCGCCCCGCGCGCGCTCCGACGCGGACCTCGGCTCCAGTCAAGCCCGCGCCG
 GCTCTCAGCTACCGCAGGAGGAGGCCACCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGAT
 GGAGGACGCGAGCAAAATTGCGAGCGCGGTGGAAGAGATGGAGGCAGAAGAGCTGCTGCTA
 AAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCAGCTATCACAATGAGCCAACACAGAC
 ACGAAGGTTGGAATAATACCATCAATGTGCACCGAGAAATCACAAGATAACCAACAACAGAC
 TGGCAAAATGGTCTTTTCAGAGACAGTTATCACAATCTGTGGGAGACGAGAAGGCAGAAGGAGCC
 ACGAGTGCAATCATCGACGAGGACTGTGGGCCAGCATGTAAGTCCAGTTTGGCAGCTTCCAGTAC
 ACCTGCCAGCATGCCGGGCGCAGAGGATGCTCTGCACCGGGACAGTGAAGTCTGTGGAGACCA
 GCTGTGTCTTGGGGTCACTGCACAAAATGGCCACCGGGCAGCAATGGGACCATCTGTGACA
 ACCAGAGGAGCTGCCAGCCGGGCTGTGCTGTGCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
 ACACCCCTGCCGTGGAGGGCGAGCTTTGCCATGACCCCGCAGCCGGCTTCTGGACCTCATCAC
 CTGGGAGCTAGAGCTGATGGAGCTTTGGACCGATGCCCTGTGTGCCAGTGGCTTCCTCTGCCAGC
 CCCACGCCACAGCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
 GAGATCTGCTGCCAGAGAGGTCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGAGGTGGC
 CCAGAGCTGGAGGACCTGGAGGAGGCTGACTGAAGAGATGGCGCTGGGGGAGCCTCGCGCTG
 CGCGCGCTGCACTGCTGGGAGGGAAGAGATTAGACTCGGACAGGCTGTGGGTAGATGTGCAA
 TAGAATAAGCTAATTTATTTCCCGAGGTGTGTGCTTTAGGCGTGGGCTGACCAAGCTTCTTCCTA
 CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGCAATTTGTCAGCT
 CCCCAGGCTGTCTCCAGGCTTCACAGTCTGGTGTGGGAGAGTCAGGCAAGGTTAACTGCA
 GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTSCCTTACCAGTTGGCAGACAGCCG
 TTTGTTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
 TGATTGGTTTTGGGGAAATGTGGAGAAGATGCCCTGCTTTGCAACAATCAACCTGGCAAAAAT
 CAACAATGAATTTCCACGCACTTCTTCCATGGGCAATAGTAAGCTGTGCCCTCAGCTGTTC
 AGATGAATGTCTGTTCACTGCAATACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
 TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATCCCTCTCAGCAGCAGCTGGG
 AGGGGGTCATTGTTCTCTCGTCCATCAGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
 CAGTCCACAGCATAGTGAAGACAGAGCACTTTCATCTGGTGTGACTCTAAGCTCAGTGCCT
 CTCCACTACCCACACAGCCTTTGGTCCACCAAAAGTGCTCCCCAAAGGAAGGAGATGGGAT
 TTTTCTTGAGGCATGCACATCTGGAATTAAGTCAAACTAATTCACATCCCTCTAAAAGTAAA
 CTACTGTTAGGAACAGCAGTGTCTCAGAGTGTGGGCGAGCCGCTCTCTAATGAAGCAATGAT
 ATTGACACTGCTCCCTCTTTGGCAGTTGCATAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA
 TACAGGTTAACTTCGACAAACAGTACTTAGGTAATTTAGGGCGAGGATTATAAATGAATTTGC
 AAAATCACTTAGCAGCAACTGAGACAAATATCAACACAGTGGAGAAATCAACCCAGCAGGGC
 TGTGTGAACAATGGTTGTAATATGCACTGCGAACCTGAACTCTAGCCACTCCACAATGATG
 TTTTCAGGTGTCAATGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCA
 CATGATTGTATAAGCATGCTTTCTTGAAGTTTAAATATGATAAACAATAGTTGCATTTAGAA
 ATCAAGTCAATAATCACTTCACTGCAAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEEAEAAAKASSEVNLANLPPSYHNETNIDTKVGNNTIHVHREIHKITNNQTGMVFSE
TVITSVGDEEGRASHECIIDEDECGPSMYCQFASFQYTCQPCRGRMLCTRDSECCGDLQCVWGHG
TKMATRGSGNGTICDNQRDCQPLCCAFQRLGFLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSMEEVRQELEDLE
RSLTEEMALGEFAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCATAAATACATCATGCAACCCAC
GGCCCACTTGTGAACCTCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATGCCAAAG
GCCTAATCAACGTTCTGTCTTCAATCTGCAAACTATATGGGGTCTGGGGCTCTTCTGGACCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTTGCCCTCCTTCTACTGGSCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAACTCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCTGACCCCTTGTGCAGATAGCCGGGTCACTTGG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGTGCTCTGGTGTCTGGAATAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAAGGTGTGTCTTGGACAAAGTCAAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGCGTGGGGTCTGTCTCTTTTCTCCGGTGCATCCCGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCTCGGGGGCTAT
GTCTACGCCAGCGGCTTCTTCAGCGTTTTCGGCATGTGTGTGGACAGCTCTTCTCTGCTTCTCT
GGAAGACCTGGAGCGGAACAAGCGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGACGAGGCGCCCGGCAACAAAGAGAGGAAGAAGTGAACAGCTCCGG
CCCTGATCCAGGACTGCACCCCAACCCCAACCGTCCAGCCATCCAACCTCACTTCGCTTACAGGT
CTCCATTTTGTGTAATAAAGGTTTATGGCCAGGCGCGGTGCTCACGCCGTGTAATCCAACACT
TTGAGAGGCTGAGCGGCGGATCACTGAGTCAGGAGTTCGAGACCAGCTTGGCCACATGGTG
AAACCTCCGCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGCATGCACCTGTCTCCCA
GCTACTCGGAGGCTGAGGCGAGGAGATCGCTTGAACCCGGAGGAGAGGTTGAGTGAAGCCGA
GATCGGCCACTGCATCCAACTCGGTGACAGACTCTGTCTCAAAACAAAAACAAACAAACAA
AAGATTTTATTAAAGATATTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NNVLALGQCVLAGAFASFYWAFHKPDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLGRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVVGGVVLSFFFFSGRIPLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYTMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGCGCCGGGCGCCGAAGCCGGGAGCCACCGCCATGSGGGGCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTCTGCTCTCGGGCTCTGCCCTGATCCTGTGCGAGCTGCTGCCCGC
CAGCCCAACTCCACCGTGAGCGCCCTCATCTTCACGTTCTTCTCTTCTGCGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAACTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGGTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGCGGCTTCTTCTTCTTCTTTTTCACCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTITTTGGTCTTTAAGTTCCTGATCTTGGTGGGCTC
ACCGTGGGTGCTTCTACATCCTTGACGGCTCCTTCACCAACATCTGTTCTACTTCGGCGTGTGGG
CTCCTTCTCTTTCATCCTCATCCAGCTGGTGTCTCATCGACTTTTGGCAGCTCTGGAACACCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCTGCTGGTACGACGGCTCTTCTTCTTCACTCTCTCTC
TTCTACTTGCTGTGATCGCGGCGGTGGCGCTGATGTTATGTACTACACTGAGCCAGCGGCTGCCA
CGAGGGCAAGTCTTCATCAGCCTCAACCTCACTTCTGTGTCTGGGTGCCATGCTGTCTGTCTG
CCAAGTCTCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCTCGGTTCATCACCTCTACACCATG
TTTGTCACTGGTCAGCCTATCCAATCCTGAACAGAAATGCAACCCATTTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCAGTGGTGGGATGCCCGAGCATG
TGGGCTCATCATCTTCTCTCTGTGACCTCTTCATCAGTCTGCGCTCCTCAGACACCCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCCACTATGCTAGACGCCACACAGCAGCAGCAGCAGAGGT
GGCAGCCTGTGAGGGCCGGGCTTTGACAAAGCAGGACGGCGTCACTACAGCTACTCTTCTTCTC
ACTTCTGCTGGTGTGCTGCTCACTGCACGTCTATGATGACGCTCACCACTGGTACAAGCCGGTGGAG
AOCGGAAGATGATCAGCAGCTGGACCGCGGTGTGGGTGAAGATCTGTGCCAGCTGGGCAAGGCTGCT
CCTCTACCTGTGGACCTGGTAGCCCCACTCTCTGCGCAACCGGACTTCAGCTGAGGCGACCTCA
CAGCCTGCCATCTGGTGCTCTCTGCCACCTGGTGCTCTCGGCTCGGTGACAGCCAACTGCCCCCTC
CCACACCAATCAGCAGGCTGAGCCCCAOCCTGCCCCAGCTCCAGGACTGCCCTGAGCCGGG
CTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCTCTTCTTCCCTCTCCTCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGTGCTGTGAGAGAGCGGGGAATCCACCACAG
TGGGGCTCCGGCACTGAAGCCTGTGTCTTCTGTCACGTCCCCCAGGGGACCTGCCCTTCTGTG
GACTTCGTGCTTACTGAGTCTCTAAGACTTTTCTAATAACAAGCCAGTGGGTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLGSAFCILSCCPASRNSTVSRLIFTFFFLGVLVSIIMLSPGVESQL
YKLPWVCEGAGIPTVLQGHIDCSLLGYRAVYRMCFATAAFFFFFFLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVSFLFILIQLVLLIDFAHSWNQRNLGKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTECPMLDATQQQQQVAAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRMISTWTAVWVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGCCAGGAACACCCGTTAAGGTGCTTCTCTTTAGGGATGGTGA
GGTGTGAAAAAGACTCCTGTAACCTCCTCCAGGATGTACCACCTGCCAGAAGACATGSGAAGC
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCTTATGAAGGAAGGAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCGTGTTGTTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTCTCTGCGAGTTTTTCGATTTAAAGTGTTAACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCACTGCCTTTTTACTAGCAANAG
TGATCCTTTTGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCGCATCATTTCATTCAATC
CTTGCTGGATTGAGACGTGGTTCCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGGGGCAGCACTTATACCTGGTGTCTTCTCTGATG
GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAAAACACTTTTAGAAGTATGTAGTACTACTTTTGTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGCAGACATAAGTTGAAATGGTGACGTC
CACTGTGTGCTTTATTGAACAGCTAATAAAGATTATTATTATTGTAATACCTCACAAAGTTGTAC
CATATCCATGCACATTAGTTGCCGTGCTGTGGCTGGTAAGGTAATGTGATGATTCACTCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGCTTGTGCTGATTTCTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTGTAGTAAGCAAGATACCTTTTATTCAATTCAAC
AGAAATGGAATTTTTTGTTCATGCTCAGATTATTATTGTATTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACATCATGCACATGTGCTCTTTGTACAGTTTTTAAAGTGTAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTCTTGTGTTGCAATTATGTGATGGCCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCAACAGACATTGTGATTATTT
TTATCATGAAATCATGTTTTTCTCTGATGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTCATATCATGTTTCTTTGGTTCAGCCAAATTCATTTAAATGAA
CTAAATTAATAA

FIGURE 14

MNHLFEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRTFCLFVTFDLLF
VTLFWIIEILNVNGGIENTLEKEVMQYDYSSYFDIFLLAVERFKVLILAYAVCRLRHHWAIALLT
AVTSAFLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVL PQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEAEKQDSEKPLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 16

MCSRVPLLLPLLLLLALGPGVQCPCGQCQCPQTVFCTARQGTTPRDPVFPDTPVGLVVFENGIT
 MLDAGSFAGLPGLQLLDLSQNIASLPQVFOPLANLSNLDLTANRLHEITNETFRGLRRLERLY
 LGKNRIIRHIQPGAFDTLDRLLLEKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDTANVE
 ALRLAGLGLQQLEGLFSRLRNHLDLVDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
 AGLAALQELDVSNLSLQALPGDLGLFPRLRLAAARNPFCVCLSWFGPWVRESHVTLASPEE
 TRCHFPKNAGRLLLELDYADFGCPATTTTATVPTTRPFVVREPTALSSSLAPTNLSPATAPATEAP
 SPPSTAPPTVGPVPQDQCPFPSTCLNGGTCHLGRHHLACLCEGFTGLYCESQMGQGTREPSPTP
 VTPRPPRSLTLGIEPVSPSTSLAVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASLAEY
 TVTQLRPNATYSVCVMPLGPGRVPEGEACGEAHTPPAVHSNHAPVTQAREGNLPLLAPALAAV
 LLAALAAGVAAVCVRRGRAMAAAAQDKGVGPGAGPLELEGVKVPLEFGPKATEGGGEALPSGSE
 CEVPLMGFPGLQSPHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGGCAGGCGGGTGGTGGCTGAGTCCGTGGTGCAGAGGCGAAGGCGACAGCTCATGCG
 GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGCGCTCGGCGTCTCTCGG
 ATGAAGAAGGCGACGAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
 AAGGACCATACTACTGCAGGCAGAGTAGTTGTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
 ATTAGAATCCTCTATTCAAGAAAGAGGAAGACAGCCTCAAGAGCCAAAGAGGGGAAAGTGTACAG
 AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
 CGGAACACAGCTTGACCGCCATTGAAGGCACAGCATGGGGAGCCCTGCCACTTCCCTTTTCT
 TTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
 CTACAACCTATGACTACAAGCAGATGAAAAGTGGGGCTTTGTGAACTGAAGAAGAGGCTGCT
 AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAATCGAATGAAAATCCTTAATGGAAG
 CAATAAGAAAAGCCAAAAAGAGAAAGCATATCGGTATCTCCAAAGGCAGCAAGCATGAACCATA
 CCAAAGCCCTGGAGAGAGTGTCATATGCTCTTTTATTGGTGATTACTTGCCACAGAATATCCAG
 GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTG
 CTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAGGCTCTTTGTATATATACAT
 TTGGAGCTCTTGGGGCAATCTAATAGCCACATGGITTTGGTAAGTAGACTTTAGTGGAAGGCT
 AATAATATTAACATCAGAAGAATTTGGGTTTATAGCGGCCACAACCTTTTTCAGCTTTTCATGATC
 CAGATTTGCTTGATTAAAGACCAATATTCAGTTGAACCTCCTCAAATTCCTGTTAATGGATAT
 AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG
 TTTGGCTGATGCCCCATAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
 GAATTTGCAGATCATTTTTACATTAGATTATCATAAATTTTAAAAATTTTCTTTAGTTTTC
 AAAATTTTGAATGGTGGCTATAGAAAAACAACATGAATATTTATACAATATTTTGCAACAATGC
 CCTAAGAATTGTTAAATTCATGGAGTTATTGTGCAGATGACTCCAGAGAGCTCACTTTCTG
 TTTTTACTTTTTCATGATTGGCTGCTCTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
 GCGTGTCTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTACTTTTCTTTGCTAATTTGG
 AAGATTAACTCATTTTTTAATAAATATTATGCTAAGATTAAAAA
 AA

FIGURE 18

MEVRIGLTLLCAVLLSLASASSDEEGSQDESLSKTTLTSDSESVKDHTTAGRVVAGQIFLDSESESL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPNKDYEEPKKVRKPALTAIEGTAHGEFCHFFFLDLK
EYDECTSDGREDRGLWCATTYDYKADKMGFCETEEEAARRQMQBAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFELKTEEGSPKGTALGFYASGLGVN
SSQAKALVYVYTFGALGGLNIAMNVLVSR

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCATTCTGCAGTGGAAATTCATGAACAGCAAGAGGACACCATTCTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAANGGAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGCCAGCCAGAACTTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGT
CCCGCGTGTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAAATGTCTCCAGTGTGGAGGTGCGCTTGCAATCGTTGGAGGGGCTATATCCATCCAAA
TATGCAGTGGAAAGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCAGCTCTC
ATGCATTGAACCAAGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAACACAACATATGGAGAAGGTTACATTGAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGTAGAGTG
CATGGACCAAGCTCTAACAAAGTCTCTTCCCTAAGACTCATTATGCCGCTGAAAAAGATGCCAAAA
TTTTCTGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGCAGTGTGACTCAGCTAACACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAGAAGACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTGGATTGCAAAAGGAGTCCCACCATCGCTGGTGGTATCCAGGGT
CCCTGCTCAAGTTTCTTTGAAAAGGAGGGCTGGAATGTTACATCACATAGGCAAGTCTGCCCT
GTATTAGGCTTTGCCCTGCTTGGTGTGATGTAAGGAAATGAAAGACTTGGCCCATCAAAATGA
TCTTTTACCGTGGCTGCCCATGCTTATGTTCCCGAGCATTTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATAAAAGATAAGTCAACCCAAAAAAGAAAAAAGAAAAA
AAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINAGVPGVLAPTDLWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
APGVHVSCIEPGLFKTNLADPVKVEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLGKNSYVNMMD
LSPVVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLKQKAEIANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGCGGTAGCATGAGGCGGAGAGTACGTCGGCGGTGCTCGGGCTTTGTGCTCGGGC
 CACTCGCTTCCAGCACCTCAACACGGACTCGGACACGGAAGTTTCTCTTTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAATGGATGATGTTGAAGTTGTTATACAATTGA
 CATTGAGAAATATATCCATGCTATCAGCTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAATATTATCAATGTCAAAAAGAATGGGTAGGTTGGTACAAATCCGT
 CGTCATTAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATTT
 TTCAAAACCAAGACCTTGTTTTCTGCTATTAAACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCCCTATATAAACCTCAAAAGGACTTTTTACAGGGTACCTTTAGTGGTT
 GCCAATCTGGGCATGCTGAACRACCTGGGTATATAAACTGTATCAGGTTCTGTATGTCCACTGG
 TTTTAGCGGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCTTAAAGGAGG
 TACATAAGATAAATGAATGTATGCTTCATTACRAGAGGAATTAAGAGTATATGCAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAAGTAGTAAAGGATGTAAACAGATTAAACAGGAAATTTGA
 GAAAAGGAGAGGAGCAGATTGAGGAGCAGAGAGAAGAACATCCAAAAGACCTTCAGGAGA
 ACATTTTTCTTTGTGAGGATACGAGACCTTTTTCCAAATCTGAATTTCTTCATTGATGTT
 ATGTCTTTAAAAATAGACATGTTCTAAAAGTAGCTGAATACCAACCACATCTGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTCGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATCAAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAGCAAACTAGGTAGTAGTAACCAAGATAAAGCATCCAAAT
 GAGCAGCCAGAAAAGATGAAGAAATGAAAAGATGAAGGGTTTGGTGAAATATCACGGTCTC
 CTACATTTGATCCTTTAACCTTACAAGGAGATTTTTATTGCTGATGGGTAAAGCAAAAC
 ATTTCTATTGTTTTACTATGTTGAGCTACTGCGAGTAAGTTCATTTGTTTTTACTATGTTCAAC
 TGTTTGCGATATACACAGATAACTCTTAGTGCTTTACTTCACAAAGTACTTTTTCAAAATCA
 GATGCTTTTATTCCAAACCTTTTTTTCACCTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTTCTTAGAATTGGAAGGTGAGACAGGCAAGTGGCTCACACTGTAAATCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAAGCTGGGCAACGTATT
 GAGACCATGCTATTAAAAAATAAATGGAAGAGCAAGAAATAGCCTTATTTCAAAATATGGAA
 GAAATTTATATGAAATTTATCTGAGTCATTAAATTCCTCTTAAGTGATACCTTTTTAGAGTA
 CATTTAGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
 AAAATTTAAAAAAGAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGGEVKGAEKNSITDSQMDDEVVYITIDIQKYI
PCYQLFSFYNSSGEVNEQALKILSNVKKNVVGMWKFRHSDQIMTFRELLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLHSLYKPKQGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSSAV
QTHSSKPFEEEDGSLKEVHKINEMYASLQEEKLSICKKVEDSEQAVDKLVKDVNRLAKREIEKRRGA
QIQAAAREKNIQKDPQENI FLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHLLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLLDTDQKRSKANTGSSNQDKASKMSSPET
DEEIERMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCAGACGCCGCGCGGCGAGGGCAGAGTCCAGCCGAGCCGAGTCCAGCCGAGCAGCGAGGGCAGGCCAA
 GCAGCGGGCAGCGAACGCCGCGCGGCCACACCTCTGCGGTCCCGCGCGCTGCCACCTTCCCTCTCTCCCG
 GCGTCCCGCGCTCGCGGCCAGTCACTTTCGCGGTTGCTGCGCCGCGAAACCCGAGGTCACACAGCGCGCGCTCT
 GCTTCCCTGGCGCGCGCGCTCTCACGCCCTCTCTCTCCCTGCGCCGCGCGCTTGGCACCGGGGACCGTGTGCTGA
 CCGAGGCCAGCTCTACTTTTCGCCCGCGCTCTCTCTCGCGCTGCTGCGCTCTTCCACCACTCCAACTCTCTCTCC
 TCCAGCTCCACTGCTAGTCCCGAGCTCCGCGAGCCCTCGCGCGCTGCTGCTGCTAGCGCGCTTCCCTCGCGCGCGAA
 GGTGGAAACGCGTCCGCCCGCGCGCACCAATGCGCACGTTCCGCTTGC CGCGCTTCTCTGACCTTGGCAGTGTCTC
 AGCGCCGCGCTGCTGCGCTGCGAGCTCAAGTCGAAAGTTGCTCGGAATGCGACGTCCTTACGTGTCCAAAGGCTTC
 AACAGAACGATGCCCGCTCCAGAGATCAACGCTGATCATTTGAAGATCTGTCGCCAGGTTCTACCTGCTGCTCT
 CAAGAGATGGAGGAAGTACAGCTCGCAAGTAAAGATGATTTCAAAAGTGTGGTCAAGCAACAGTGCATCATTTG
 CAAGCTGTCTTTGCTTCAAGTTAAGAAAGTTTGAATCTCTCAAAGAACTACTTGAATGAGAGAAATCCCTG
 AATGATATGTTGTGAAGACATATGGCCATTATACATGCAAAATTTGAGCTATTTAAAGATCTCTCTGTAGAGTTG
 AAACCTACTACTGTTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGCTCGCTCTCGAGCGGATG
 TTCCGCTGCTGAACCTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGAGCACTGAAG
 CCTTTCGAGAGTCTCCCTCGCAAAATGAAGCTCCAGGTTACTCGTGTCTTTGTAGCAGCGCGTACTTTCGCTCAAGGC
 TTAGCGGTTGCGGAGATGTCTGAGCAAGGTCTCGGTGTAACCCACAGCGCCAGTGTACCATGCTCTGTGAAG
 ATGATCTACTGCTCCCACTGCGCGGCTCTCGTGAAGTGTGAAGCCATGTTACAACTACTGCTCAACATCATGAGGCG
 TGTGTGCCAACCAAGGGATCTCGATTTGAATGGAACAAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
 GAGGCTCCTTCAACATTGAATCGGTATGGATCCCATGATGTGAAGATTTCTGATGCTATATGAACATCATGAGAT
 AATAATGTTCAAGTGTCTCAGAAGGTTTCCAGGATGTGGAACCCCGAGCGCTCCAGCTGAGCAATTTCTCGT
 TCCCTCTCTGAAAGTGCCTCAAGTGTCTGCTTCAAGCACATCAACCCGAGGAACGCCAACCAAGCAGCTGGCACT
 AGTTTGGAGCAGCTGGTTACTGATGTCAAGGAGAACTGAJACAGGCCAGAAATTTCTGCTCTCCCTCCGAGCAAC
 TTTTTCACAGATGAGAGGATGGCTGAGGAAACCGCAATGAGATGACTGTGGAATGGGAAAGGCAAGACAGGTATC
 CTGTTTGCAGTGAAGGAAATGATTAACCAACAGGGCAACCAACAGAGTTCAGGTTGACACCAAGCAACAGAC
 ATACTGATCTCTCTCAATCATGCTCTTCAAGTATGACCAAGAGTGAAGATGCATACAAATGGGACGACGCTG
 GACTCTTTGATATCAGTGATGAAGTAGTGAGAGGAAGTGAAGTGGCTGTGAGTATCAGAGTGCCCTTCAGAG
 TTTGACTCAATGCCACTGACCTGCTGGGAAGTGCCCAATGAGAAAGCCGACAGTGTGCTGCTGCTCTGCGGGCA
 CAGGCTCACTCTCACTGTCTCTGCTATCTGTTCTGCTGTTATGACAGAGAGTGGAGATTAATTTCTCAAACTCTGAG
 AAAAGTGTTCATCAAAAGTTAAAAGGACCAAGTTATCACTTTTCTACCATCTGATGACTTTTCTTTTAAATGAA
 TGACACCAATGTACAGTTTTACTATGTGGCCACTGTTTGAAGTGTGCTGACTTTTCTTCTCATTCAGTTTGGG
 AGGAAAGGAGCTGTGCTGATGTTGTTGTTCTGCTGCCCGCAACCATGTTAAAGTGTGCTTACAGTGTAGGATCAGAA
 CTATAGTATGTTGCAATTTGATTTTATCACTCTATTATTTGTTTGTATGTTTCTTCTCAATTTGTTGTGGGTT
 TTTTTTTCCAACTGTGATCTGCGCTTGTCTTCAAGCAAAACAGGCTCCCTTCTGGCAGTAACATGTACGTATTC
 TCTGAATATTAATAGCTGTACAGAGCAGGTTTATTTATCATGTATCTTATTAAGGAAAAAGGCCAAAAAGC

FIGURE 24

MARFGLFALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDFKSVVSEQC�HQLQAVFASRYKKFDEFFKELLENAEKSNDMFVKTYGH
LYMQNSLFFKDLFVELKRYVVGNVNLEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPFQGDVPRKLKLVTRAFAARTFAQGLAVAGDVVSKSVVNPTAQCTHALLKMIYCSHCRL
VTVKPCYNYC�NIMRGCLANQGDLDFFWNNFIDAMLVVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFGCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLQAKKFWSSLPNVNCNDERMAAGNGNEDDCWNGKGKRYLFAVTGNGLANQGNNEVQVDTIS
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAGCATAGACCACAGGCTGAGTATCCTGAC
CTGAGTCATCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGCTCTCTAGCAGCCTGAATCCAGGGGTCCGCAGAGGCCACAGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAGACACCAAGGCACCACAGAAAGCCAAACAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAAGACACTC
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTTGTGTTGCTCTCTCTAGTGTCTTCTTCTCTGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTTAACCTTAAATGCAATCAGGAAGTAGCAACAGAAGTCATATAATATTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLLPLMIAMSVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPDHFKNVKKTRHQRRHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
 AGCTGGTCTGCGCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTGTCTGCTTCTTACCCCTGCC
 CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCCCTTACCTGA
 TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAGAAACGGGAGCTCTTCAGCCAG
 ATAAAGGGGCTTACAGGAGCCTCCGGGAAGTGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
 CAACTTTAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCACCTTTGAGA
 AGTTCTTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGTTTGTGGTGGCTCCT
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGTG
 CTCGTGTGCAGAGCCCAAGGAAGGTCTCGAGGAGTCCGGAGAGTACTGAGACCGGAGGTGTGC
 TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCCTCATGTGGCAGCAAGTTTT
 GAGGCCACCTGGAACACATTGGGGATGGCTGCTGCCTCAOCAGAGAGACTTGAAGGATCTTGA
 GAACGCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
 GGCCCCACATCATGGGAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACCTATTGTCTC
 TTCCCCAGCCTCCAATTAGAACAAGCCACCCACAGCCTATCTATCTTCCACTGAGAGGGACCTA
 GCAGAAATGAGAGAGCATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGSC
 AATCTCTAACTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
 AAACACTAGGACCTGTTGTATCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
 CCAATGTTGTCCTTTCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCTGAGGCTACAC
 CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCATCATCCTGCCAAGCCCCCTGAC
 CCTCTCTCCCACTACCACCTTCTCTCTGAGCTGGGGGCACAGGGAGAAATCAGAGATGCTGGGG
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAAATATTTTTAATAAATAGACGAA
 ACCACG

FIGURE 28

MDILVPLQLQLVLLLTLP LHLMLALLGCWQPLCKSYFFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELCGGTGANFQFYPPGCRVTC LDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQS PRKVLQEVRRVLRPGGVLF FNEHVAEYPGSWAFMWQQVFPEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLVPVGHMKGAVKQSFPSKALICSPSL
QLEQATHQFIYLFPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCTATCCACCTCCCCAAGCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCTTCTTTCCCCCTTCCTTTCTGTGTCTCTGCTCAT
CGGCTGCCATGACCTGCAGCCAGCCAGCCCGTGGGGAGGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGGAGATAGGGACAGAAAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTCAATAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCAGGCTACCAGTT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAACACCTGAGTGTCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCTCCACCGAATCTAAGTGGAAATCATGTGCGG
 GAAGAGATACAATCCTTGGCCCTGTGTATCTCGCATTAGCCCTTGCTTTGGCCATGATGTTTACC
 TTCAGATTCATCACCACCTTCTGGTTACATTTTCATTTGATTGTTATTTGGGATTGTTGTT
 TGTCTCGGGTGTATTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGAGACA
 CAGAAAGGAAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCAGGCAGTG
 CTGCTCGCTTGATTTTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
 AATAAAGCCATCAGCAGTGTCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
 TTTTCTTCTGGGTCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACCTGCAGGAGCTGCCAGGTT
 ATGGAAGGCGGCCAAGTGGAAATATAAGCCCTTTCGGGCATTCGTTACATGTGTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAAATGACTATAGCTGGGGCAG
 TGCTTACTTGTATTTCAACAGAAATAAAATGATCCTCTGATCATCCATCCTTTGCTCTCTC
 TCCATTTCTCTTCTTACCATCAAGGAACCGTTGTGAAGGGTCATTTTAATCTCTGTGGTGAG
 GATTCAGAGAATCATTGTGATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCAATGT
 CCAGGTACTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTCAACTGCTATTATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCCAGAAGTCAAGTCATTTACATCTAATTAACCTGCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTAACTAC
 AATCGGGCATTCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTATCTGTGTTTGAAACTGTGCTGGATGCATTTTCTGTGTTTGTGCTGTTGATC
 TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
 GGGACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAACATTT
 TCCTTCTAAGAGCCATTTACAGAAATAGAAATGAGACCCTAGAGAAAAAGTTAGTGAATTTTTTT
 TTAAAGACCTAATAAACCTATTCTTCTCTCAAAA

FIGURE 32

MSGRDTILGLCILALALS LAMMFTFRFITLLVHIFISLVILGLLFVCGVLNWLYYDYTNLSIE
LDERENMKCVLGF AIVSTGITAVLLVLI FVLKRRIKLTVELFQITNKAISSAPFLFQPLWTF A
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDDPPDHPILSSLSILFFYHQGT VVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFOVWAVPLLVAFFAYLVASFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 34

MRTVVLTMKASVIEFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPPTLPRPSPSAASTTSIPRPQSVGHRSEQMDLWSTATYTSSQNRPRADPGIQRQDPSGAAAFQKP
VGADVSLGLVPKEELSTQSLEFVSLGDPNCKIDLSLFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHTNLKTHNSRDLKTAIEKIQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVMVDGWPTDKVEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVLRVCDTDRACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQTYEQRLFEFGDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSHFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCAGGAGATTGCGTTCGGTTAGGAGGTGGCTGGTTGTGGGAAAAGCTATCAAGGAAGAAATTC
 CAAACCATGTCTTTTTTCTGTTCAGAGTAGTTCACACACAGATCTGAGTGTTTAATTAAGCATGGAAT
 ACAGAAAACACAAAAAATTAAGCTTTAATTTCACTCTGGAAATCCACAGTTTTCTTAGCTCCCTGGAGCC
 GGTTGACCTGTTGGCTCTCCCGCTGGCTGCTCTATCACCTGGTGCTCTCCGACTACTCAOCCGAGTGTA
 AAGAACCCTCGGCTCGCGTCTCTGAGCTGCTGTGAATGGCTCGGCTCTCTGGACTGTCTCTCCGAGTA
 GGATGTCACTGAGATCCCTCAAATGGAGCCCTCTGCTGTCACTCTCGAGTTCTTTGTGATGTGGTA
 CTCAGCCTCCOCCACTACAATGTGATAGAACCGCTGAACCTGGATGTACTTCTATGAGTATGAGCCGATTTA
 CAGACAAGACTTTCACCTTCACACTTCGAGAGCAITCAARCTGCTCTCATCAAAATCATTCTCGTCACTC
 TGGTGACCTCCCACCTTCAGATGTGAAGCCAGGCGCAITAGAGTTACTTGGGGTGAANAAGTCT
 TGGTGGGGATATGAGGTTCTTACATTTTCTTATTAGGCCAAGAGGCTGAAGAAGCAAAATGTGGC
 ATGTCTCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
 ACCTGACCTGAAAACCAATTATGGCAITTCAGGTGGTAACTAGTATTTGCCCAATGCCAAGTACGTAATG
 AAGACGACACTGATGTTTTCACTCAATACCTGGCAATTTAGTGAAGTATCTTTAAACCTAAACACTCAGA
 GAAGTTTTTACAGGTTATCTCTAATTGATAATTATTCTATAGAGGATTTACCAAAAACCCATATTT
 CTTACCAAGAGTATCCTTCAAGGTGTTCCCTCCATCTGACAGTGGGTGGGTATATAAGTCCAGAGAT
 TTGGTGCCAGGATCTATGAATGATGGGTCACTAARACCCATCAAGTTTGAAGATGTTTATGTCCGGAT
 CTGTTTGAATTTATTAAGTGAACTTCATATTCCAGAAGACACAAATCTTTCTTCTATATAGMATCC
 ATTTGGATGTCTGTCACTGAGAGCTGTGATGTGACCCATGGCTTTTCTTCAAGGAGATCATCACTTTT
 TGGCAGTCTATGCTAAGGAACACCCATGCCATTATTAATTTTCACTTCACTTCTACAAAAGCCTAGAAGGACAG
 GATACCTTGTGGAAAGTGTAAATAAAGTAGTACTGTGGAAAATTCATGGGAGGTCAGTGTCTGGCTT
 ACACGAACTGAACTCATGAAAAACCCAGACTGGAGCTGGAGGGTTACACTTGTGATTTATTAGTCAGG
 CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATGGAGGTTTTGCTAAGAAATTAATAGG
 ACCAAACAATTTGGACATGTCACTTCTGTAGACTAGAAITTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
 CTAGGCTGAAAAACAAAACAACTGTAGAGTTTTATTTATTGAACAACTGTAGTCACTTGAAGGTTTTGTGTA
 TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAATCTTCTCACTGAAGTTATA
 CTGAACAAAATTTACCTGTTTTTGTCTATTTATAAGTACTCAAGATGTTGAGTATTTCAAGTTATT
 ATTAATTTAAATTACTTCAACTTTGTGTTTTTAAATGTTTGAAGATTTCAATACAGATAAAAAGGATAG
 TGAATCATCTTTACATGCAACATTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
 TTAATGTAAGTCATAGGTCATTTATGTCATATCAGTAACTCTCTGGACCTTGTAAATATTTTATCTGTGGT
 AATATAGAGAAGAAATTAAGCAGAAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNMWYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHEPSDVKARQAIRVTWGEKSWWGYEVLTFLLGQEAEEKDKMLA
LSLEDEHLLYGDIIRQDFLDITYNNLTLTIMAFRWVTEFCPNAKYVMKTDTFINTGNLVKYL
NLNHSEKFFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVIGICLNLKVNIIHIPEDTNLFFLYRIHLDDVQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

C8CTCGGGCACCAGCCGCGGCAAGGATGAGCTGGGTGGCTGGACGAGTTGGGGCTCAGTTTTCCTCAGCTCCTTCATC
 TCGTCTTCCCAAGAGATACACAGTCATTATGAAGGCTGCCCCGGACGAGGTGAATATCATGTCTGGGAGTGGCTGTG
 AATATGATCAGATTGAGTGGCTGCCCCGGAAGAGGGAGCTGGGGTTATACCATCCCTTGTGCGAGGATGAGGAGAA
 TGAGTGTGACTCCCTGCTGATCCACCCAGTTGTACCATCTTTGAAACATGCAAGAGCTGGCGAATGCTCATGTGGGGGGT
 ACCTTGGATGACTCTATGTGAAGGGGTTCTACTGTGAGAGTCCGAGCAGGGCTGGTACGGAGGAGACTGCATCGGATGTG
 GCGAGGTTCTGCGAGCCCAAGGGTCAGATTGTGTGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCAITCATGC
 TAAACCTGGGTTTGTATCCAACTAAGATTGTGATGTGTGAGTGGAGTTTGCATGATGTGCCAGTATGACTATGTGTAG
 GTTCTGTGATGAGACACCCGCGATGCCGATCATCAAGCTGTCTGTGGCAACGAGCGGCCAGCTCCCTATCCAGAGCATAG
 GATCCCTCACTCCAGCTCCTTCCACTCGATGCTCCAGAAATTTGAGGGTTTCCATGCCAATTATGAGGAGATCAAGC
 ATGCTCCTCATCCCTTGTTTGATGACGGCACGTGGCTCTTGACAAGGCTGGATCTACAGTGTGCTGTTGGCAGG
 TATACTGGGCGAGCGCTGTGAATATCTCTGAGAAAGAACTGCTGACACCTTGGGGGCCAGTCAATGGGTACAGAA
 TAAACAGGGGGCCCTGGGCTTATCAAGGACGCCATGCTAAATTTGGCACCGTGGTCTTCTTCTTTGTAACTCATCTATGT
 TCTTAGTGGCAATGAGAAAGAACTTCCACAGAAATGGAGAGTGGTCAGGGAAACGCCCATCTGCATAAAGGCTGCCGA
 GAACCAAGATTTCAAGCTGTGAGAAAGAGAGTTCTCCGATGCGAGTTCAGTCAAGGGAGACACCATACACCACTAT
 ACTCAGGGGCTTCAGCAGCAGAACTGCGAGATGCCCTTACCAAGAACCCAGCTTCCCTTTGGAGATCTGCCCATGGG
 ATACCAACATCTGCAACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCCGCGCTGGGAGCAGCAGAGGAGCATGT
 CTGAGGACTGGGAGTGGAGTGGGCGGGCACCATCTGATCCCTATCTGGGGGAAATGAGACATCACTGCTGCCAAGA
 CCGAAGGTTGGCTGGCGCTGGACGGCACCATCTACAGAGAGCCAGCGGGTGCATGCGCAGCTTACACAGAGGAGC
 GTGGTCTCTAGTCTGCGAGGCTGCCCTGGTGAATGAGCGCACTGGTGGTGGTGGCCACTGTGTTACTGACCTGGGAG
 GTCACTGATGATCAGACAGCAGCTGAAAGTTGTTTTGGGGAAATTCACCGGATGATGACCGGGATGAGAGACCATCC
 AGAGCGACAGATTCTGCTATCATCTGATCCCACTATGACCCCATCTGCTTGAATCTGACATCGCCATCTGAGAGCT
 CCTAGACAGGGCCCTATCAGACCCCGAGTCCAGCCCATCTGCTGCTGCGAGTGGAGTCTCAGCATCTCCTTCCAGAG
 TCCCATCATCTGTGGCTGGCTGGATGTCTTGGCAGCATGAGGAGCTTGGCTTCAAGAACGACCACTGGCTCTGGGG
 TGGTCACTGTGGACTGCTGCTGTGTGAGGAGCAGATGAGGAGCATGGCATCCAGTGGAGTGTCTAGTATACATATT
 CTGTGCGAGCTGGGAAACCACTGCCCTTCTGATATCTGCACTGAGAGCAGGAGCATCGCGCTGTGCTTCCGGGGA
 CGAGCATCTCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAACAATGAGCCACAGGCTCTCA
 CTGCTTCCACAGGTGCTGGCTTTTAAAGACTGGATTAAAGAAATATGAATGAGACATGCTATGCACTCCTTGAAGA
 TGTTCCTGTATCCCTCTGTACGTGTGATGCTGAGGAGTGGGGCTGAATGTGATTTGGCTGTGACTTGGCT
 GTGCGAGGCTTCTGACTTCAGGAGCAAACTCAGTGAAGGGTGAAGTGAAGCTCAATGGTGGTAGGCTGATGCCGCTCCA
 CTACTAGGACACCAATTGGAGATGCCAGGGCTGCAAGAAATAGTTTCTTCAAGAGACCATATACAAACCTCTCCA
 CTCACCTGAGCTGGTGGTCTTCCGCCAATCTTCAATACGAAATGGCATCAGCTTGACCAAGGAGATCTGGGCTCATAG
 GCGCCCTTTGAGGCTCTCAGTTCTAGAGAGCTGCTGTGGGACAGCCAGGGCAGCAGAGTGGGATGTGTGATGCTT
 TGTGTACATGGCCAGCTACAGCTTGGCTCTTCTCTCCCATCTCTTGTACATACATTTAATAAATAAGGGTTGGCTCT
 GATCTACAAAAA
 AA
 AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENWIMCRECEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCABCRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMSLEFDYMCQYDYVEVRDGNRDGQII
KRVCGNERPAPIQSIGSSLHVLPHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE
KRTCQNGEWSGKQPIKICAKREPKISDLVRRRVLPQVQSRETPLHQLYSAFQKQLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRLGSSRRRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPNQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFECA
SWEPTAPSDICTAETGGIAAVSFPGRASPEFRWHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTCTCTACATCTCTCATCTGAGAATCAGAGAGCATAACTCTTACGGGCGCGTGAATTTATTACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCGCTTGAGCTCAGAGAGGTGGCCAGGAGATGCAGCACTGCTCGAGAGATAGGGCGCTCTGTGTC
TGGCTTGCCCTTGGCTCAGTCTGCTAACTACATTGACAAATGGGCAAGCTGGACTTCTGTATTACAGAACTGTGTA
AAGGTGCTTCCACTACGGCTGACCAAGATAGGAAGGGCGCTCACAAGATGGCTGCCAGCGGCTGTGCGAGCC
TCACAGCCACGGCTCTCTCCACAGAGTTTCTGCAGTGCACCATCTCTTAATGACAGACGAGCTGGCTAGACA
ACCTCGCTACGTGTCTCGGAGAGGAGGGGAGCGAGCAGCARTCAGCCGAGTGGACTTGGCGGAGCAACGAACTA
GGGACGGGCGCTTTGAGAGATCCACTATTAGAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
GGACAAGAGCGGAGTGCAGTTGCCAACCATGCCAGCAGGGCAGGGAATTTCTGAAAACACCACTGCCCTGAG
TCTTTCCAGGTTGTACCACTGATTCAGATGGTGAATTACAGCATCAAGATCAATCGAGTAGTCCAGTGA
GCTCTCTATTAGGCTGGTGGAGGTAGCGAAACCCCACTGGTCCNATCATTTACCAACACATTTATCGTAGGGG
TGATGCCCAGAGACGGCGGCTACTGCCAGGAGACATCATTTAAAGGTCAACGGATGGACATCAGCAATGCCCTC
ACAACTACGTGTGCTCTCTCGGCGACCCCTGCCAGGTGCTGTGGCTGACTGTGATGCTGAACAGAGTTCCGCA
GCAGGAACATGGAGAGGCGCCGGATGCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAATGAGC
CGAGGAGCAGCTTGGATAAAACTGGTGGCAAGGTGGATGAGCTGGGGTTTTCATCTTCAATGTGTGATGGG
GTGTGGCATATGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATTGGCA
GCCCGAAGTGGCGCTCATCTGATTCAGGCCAGTGAAGAAGTGTACCTCTGCTGTCCCGCAGGTTCCGAGC
GAGCGCTGACATCTTTCAGGAAGCGGCTGGAACACCAATGGCAGTGGTCCCGAGGCGAGGAGAGCAACA
CTCCGAGCCCTCCATCTCAATTACTTTGTCAATGAGAGGGGTGAATTCGAAAGACCCCGGTGAATCTCTCG
GCATGACCGTGGCAGGGGAGCATCACATAGAGAAATGGATTTGCTATCTATGTCACTAGTGTGGAGCCGAGAGG
TCATAAGCAGAGATGGAGAATAAAAAACAGTGACATTTTGTGAATGTGGATGGGTGCACTGACAGAGGTGAGCC
GGAGTGGGCGAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAGCTTTGGAAGTCAAGAGTATGAGC
CCAGGAAGACTGCAGCGCCAGCAGCGCTTGGACTCCACACCAACATGGGCCACCCAGTACTGGTGGCCATCTCT
GGGTCTATGTGGTGGATTAACACGGGTGCTTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGT
TGGGCTTCTGCAATGTAGGAGGTTATGAAGAATACAAATGGAACAAACCTTTTTTCATCAATCCATTGTTGAAGGA
CACGAGCATACAATGATGGAAGAAATAGATGTGTGATTTCTTCTGCTGCAATGGTAGAATACATCAGGAATGA
TACATGCTTGCTTGGCAAGCTCTGAAGAACTTAAAGGAAGATTAATCTAATTTGTTCTTGGCTGGCCTG
TTTATAGATCAATGATGGGTGAGGAGAAACAGAAAAATCACAATAGGCTAAGAGTTGAACACTATATTATTC
TTGTGATTTTATATTTAAAGAAAGATACATTGAAAAATGTGAGGAAAGATGATCATCTAATGAAGCGAGTT
ACACCTCAGAAAAATATGTTCCAAAAAATTAACACTACTAGTTTTCCTCAGTGGGAGGATTTTCATTACTCTAC
AACATGTTTATATTTTCTATCAATAAAAGCCCTAAACAACTAAATGATTGTTGTATACCCACTGAAT
CAGCTGATTTAAATTTAAATTTGTTATATGCTGAGTCTGCCAGGGTACATATGCCATTTTAAATTTACAGCT
AAAAATTTTTTAAATGCAATGCTGAGAAAGCTTGCTTTCATCAACCAAGATATAATTTTTGAGAGTAA

FIGURE 40

MKALLLLVLPWLSFANYIDNVGNLHFLYSELCKGASHYGLTKDRKRSQDGCPCDGCASLTATAPS
PEVSGAAATISLMTDEPLDNFAYVSSAEDGQPAISFVDSGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIILKVNQMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVLNKSSEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDIFQEAGWNNSGNSWSPG
PGRSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPYIVISVEPGGVISRDGR
IKTGDILLNVGDVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNNHMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRNTAGSLGFCIVGGYEEYNGNKPFPIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTGCATCAAGTTCCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
 CTTGGCTGCAGTGAAGCAGAGAGATAGATATTATTACAGTAATAAAAAACATGGGGCTTCAACTGACT
 TTCCACCTTTCTTACAAATCCGATTACTGTTGCTGTGACTTTTGGCTGACAGTGGTTGGTGGGC
 CACCAGTAACACTCTCGTGGGTGCCATTCAAGAGATTCTTAAGCAAGAGGAGTTCATGGCTAATTTC
 ATAAGACCCCTCATTTTGGGAAGGGAAAAAAGCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
 GACAACTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAGCTCATTTTCAAAACAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGGTCGCCATCTCTGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG
 TAAAAAGTTTAATCGAGCCAAACTCTTGAATGTGGCTATCTAGAAGCCCTCAAGGAAGAAAAATTGGG
 ACTGCTTTATATCCACGATGTGGAACTGGTACCCGAGAATGACTTTAACTTTTACAAGTGTAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
 GGGTGTACTGCCCTAAGCAGAGAGCAGTTTTCAGGTGAATGGATTCTTAACAACTACTGGGGAT
 GGGGAGGCCAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAGAATGAAAATTTCCCGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAAGACG
 GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAAACAAATCCTTTATATATCAACATCAAGTGGATTCTGGTTTGGTGCAATGACCC
 TGGATCTTTTGGTGATGTTTGGAAAGACTGATTCTTTGTTTGAATATTTTGGCTAGAGACTTCAA
 ATAGTAGCACACATTAAAGAACTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAAACAGACAGCTTCTTAGTCATTTTGAT
 CATGAGGGTTAAATATTGTAAATGGTACTTGAAGGACTTTATATAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTATTTAAATTTGAAGTAATATATATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTGATTTATCTGTACAATCATCTGTGAAGTGGTGGTGAGGT
 GAGAAGGGCTGCACAAAAGGGGAGAAAGGGCAGCAATCAGGACACAGTGAATTTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGCAGCAGTAGCTGAGCTGTTGCAAGGTGCTGATAGC
 CTTCAAGGGAGGACCTGCCAGGTATGCTTCCAGTGTGATGCCACAGAGAAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTTAATATAATAAATATGTCTATCAAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLT FHL SYKFRLLLLLTLC LTVVGWATSN YFVGAIQE I PKAKE FMANFHK T LILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEE VQAENPKVSRGRYPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYG IYVIHQAE GKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLKCEEHPKHLVVGRNSTGYLRYSGYFGGV TALSREQFFKVN GFSNNYWG WGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCAGATCTGGG
 CCGCTTGCCCTCTGCTCCTCCTCCTCGCCAGCCTGACCACTGGCTCTGTTTCCCAACAG
 ACGGGCAACTTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
 GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
 ATCGATCAAACTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCGTCCCTCCC
 TTCTTATTTATTCTGCTGCCCGAGAACATAGGTCTTGAATAAAATGGCTGGTCTTTTGT
 TCCAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0
 1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 44

MALSSQIWAACLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GFGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGUTGGTTCCCCAACATGCCTCACCC
 TCATCTATATCCTTTGGCAGTTCACAGGGTCAGCAGCCTCTGGACCCGTGAAGAGAGCTGGTCGGT
 TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
 GACCTTCAACACAAACCCCTCTTGTACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
 ATCGTAATAGGGAGAGAGTAGAUTCCTCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
 AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
 GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
 ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAAGGATGTGATT
 TATACCTGGAAGGCCCTGGGGCAAGCAGCCAAATGAGTCCCATATGGGTCCATCCTCCCCATCTC
 CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCTGTGCAGCAAACT
 TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCTCATG
 GTCTCTCTGTGTCCTGTTGGTGCCTTCCGTGCTCAGTCTCTTTGTACTGGGGCTATTCTTTG
 GTTCTGAAGAGAGAGAGACAAGAAGGTACATTGAAGAGAAGAAGAGATGGACATTGTCTGGG
 AAATCTCTAACAATATGCCCCCATCTGGAAGAAACACAGAGTACGACACAATCCCTCACTAAT
 AGACAATCCTAAAGGAAGATCCAGCAAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
 GGAAAAATCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCTATGAAGATGTTA
 TCTAGACAGCAGTGCACTCCCCATAAGTCTCTGCTCA

FIGURE 46

MAGSPCTCLTIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVVTFNTTPLVTIQP
EGGTIIIVTQNFRNRERVDFFDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGOAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTGYDTIPHINRTILKEDPANTVYSTVEIPKMKMENPHSLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGSATGGACATCCTGCAATGG
 ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTCTCAATGCGATACCTCTAATTGTCAGCT
 TAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCAGGAATT
 ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
 CTGCAACAACAGAACTGGAATGTTTCTTTTCATCATTTTTCAGTGTGATCACAGTCATTGGTGCTC
 TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
 AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
 CTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCACCAGTAACG
 ACACCATGSGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAAACAACATAGG
 CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAATTCGAGGTCCTGTTTGGGCT
 CAGTCAGTAGTCATCGSTTTCCCTGGCTGTCTGTGTGGATCTCTAAGCGAAGAAGTCARAATTG
 TGTAGTTTAATGGGAATAAATGTAAGTATCAGTAGTTTGAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
21

FIGURE 48

MTCCEGWTSCNGFSLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFFGIIGAGIMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNPSNSNANCEFSL
KNISDIHPESFNLOWFFNDSCAPPTGFNKPTSDNTMASGWRASSPHFDSEENKHLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCTCTCCCCCTGCAGCCCTGCCCTCGAAGCTGTGA
CATCGAGAGAGTGACCCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAAGCCAATGACC
CATTTGCCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAATACAAGAG
CAGCCAGAAGCAGCACAGTCTGTACCTGAGAAGGCCATCCCCTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCGGGAGGCCTTATCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACACCCCCCTCA

country	city	address	phone	fax	telex	internet	url
USA	San Francisco	1000 California St., Suite 100	415 774 2200	415 774 2201		www.elsevier.com	
USA	San Diego	3600 La Jolla Village Dr., Suite 100	619 594 9000	619 594 9001		www.elsevier.com	
USA	Washington, DC	1000 17th St., NW, Suite 100	202 326 7000	202 326 7001		www.elsevier.com	
USA	Atlanta	1000 Peachtree St., NE, Suite 100	404 525 7000	404 525 7001		www.elsevier.com	
USA	Chicago	1000 N. Dearborn St., Suite 100	312 527 7000	312 527 7001		www.elsevier.com	
USA	Los Angeles	1000 Wilshire Blvd., Suite 100	213 551 7000	213 551 7001		www.elsevier.com	
USA	New York	1000 10th Ave., Suite 100	212 512 7000	212 512 7001		www.elsevier.com	
USA	Philadelphia	1000 Locust St., Suite 100	215 527 7000	215 527 7001		www.elsevier.com	
USA	San Jose	1000 N. Zeeb Rd., Suite 100	408 531 7000	408 531 7001		www.elsevier.com	
USA	Seattle	1000 1st Ave., Suite 100	206 461 7000	206 461 7001		www.elsevier.com	
USA	Portland	1000 N. Broadway, Suite 100	503 461 7000	503 461 7001		www.elsevier.com	
USA	Denver	1000 17th St., Suite 100	303 527 7000	303 527 7001		www.elsevier.com	
USA	Phoenix	1000 N. Central Ave., Suite 100	602 527 7000	602 527 7001		www.elsevier.com	
USA	San Antonio	1000 N. N. St., Suite 100	214 527 7000	214 527 7001		www.elsevier.com	
USA	San Diego	1000 N. La Jolla Village Dr., Suite 100	619 594 9000	619 594 9001		www.elsevier.com	
USA	San Jose	1000 N. Zeeb Rd., Suite 100	408 531 7000	408 531 7001		www.elsevier.com	
USA	Seattle	1000 1st Ave., Suite 100	206 461 7000	206 461 7001		www.elsevier.com	
USA	Portland	1000 N. Broadway, Suite 100	503 461 7000	503 461 7001		www.elsevier.com	
USA	Denver	1000 17th St., Suite 100	303 527 7000	303 527 7001		www.elsevier.com	
USA	Phoenix	1000 N. Central Ave., Suite 100	602 527 7000	602 527 7001		www.elsevier.com	
USA	San Antonio	1000 N. N. St., Suite 100	214 527 7000	214 527 7001		www.elsevier.com	
USA	San Diego	1000 N. La Jolla Village Dr., Suite 100	619 594 9000	619 594 9001		www.elsevier.com	
USA	San Jose	1000 N. Zeeb Rd., Suite 100	408 531 7000	408 531 7001		www.elsevier.com	
USA	Seattle	1000 1st Ave., Suite 100	206 461 7000	206 461 7001		www.elsevier.com	
USA	Portland	1000 N. Broadway, Suite 100	503 461 7000	503 461 7001		www.elsevier.com	
USA	Denver	1000 17th St., Suite 100	303 527 7000	303 527 7001		www.elsevier.com	
USA	Phoenix	1000 N. Central Ave., Suite 100	602 527 7000	602 527 7001		www.elsevier.com	
USA	San Antonio	1000 N. N. St., Suite 100	214 527 7000	214 527 7001		www.elsevier.com	
USA	San Diego	1000 N. La Jolla Village Dr., Suite 100	619 594 9000	619 594 9001		www.elsevier.com	
USA	San Jose	1000 N. Zeeb Rd., Suite 100	408 531 7000	408 531 7001		www.elsevier.com	
USA	Seattle	1000 1st Ave., Suite 100	206 461 7000	206 461 7001		www.elsevier.com	
USA	Portland	1000 N. Broadway, Suite 100	503 461 7000	503 461 7001		www.elsevier.com	
USA	Denver	1000 17th St., Suite 100	303 527 7000	303 527 7001		www.elsevier.com	
USA	Phoenix	1000 N. Central Ave., Suite 100	602 527 7000	602 527 7001		www.elsevier.com	
USA	San Antonio	1000 N. N. St., Suite 100	214 527 7000	214 527 7001		www.elsevier.com	
USA	San Diego	1000 N. La Jolla Village Dr., Suite 100	619 594 9000	619 594 9001		www.elsevier.com	
USA	San Jose	1000 N. Zeeb Rd., Suite 100	408 531 7000	408 531 7001		www.elsevier.com	
USA	Seattle	1000 1st Ave., Suite 100	206 461 7000	206 461 7001		www.elsevier.com	
USA	Portland	1000 N. Broadway, Suite 100	503 461 7000	503 461 7001		www.elsevier.com	
USA	Denver	1000 17th St., Suite 100	303 527 7000	303 527 7001		www.elsevier.com	
USA	Phoenix	1000 N. Central Ave., Suite 100	602 527 70				

Important features:

amino acids 1-16

amino acids 36-59

amino acids 41-47, 45-51, 84-90

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGGAGGG
 AGGACAGGGAGTCSGAAGGAGGAGGACAGAGGAGGGCCACAGAGACGCAGAGCAAGGGCGGCAAGG
 AGGAGACCTTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
 GGGCCCCCTGGCTGCCTCCTGCTGGCCCTCTGGCTGGGCAGTGGGAGGCTGGCCCCCTGCAGAG
 CGGAGAGGAAAGCACTGGGACAAATATTGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
 GCGAAGGGGTGGGAAGGCCATTGGCAAGAGGCCGGAGGGGCAGCTGGCTCTAAGTCAGTGAG
 GCCCTTGGCCAAGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
 AGCAGATGCTTTGGGCAACAGSGTGGGGGAAGCAGCCCATGCTCTGGGAACACTGGGCACGAGA
 TTGGCAGACAGGCAGAAGATGTCAATTCGACACGGAGCAGATGCTGTCCGCGGCTCTGGCAGGGG
 GTGCCITGGCCACAGTGGTGTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGTGG
 CCTTGGAGGCCAGGGCCAGGCAATCCTGGAGTCTGGGGATCCGTGGGTCCACGGATACCCCG
 GAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
 GGGCCACCAACTTTGGGACCAACTCAGGAGCTGTGGCCAGCCTGGCTATGGTTCAGTGAG
 AGCCAGCAACAGAAATGAAGGGTGACGAATCCCCACCATCTGGCTCAGTGGAGGCTCCAGCA
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAA
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGCAGCAG
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGCAGTGAGT
 CCTCTGGGGATCCAGCACCGGCTCCTCCTCCGSCAACACAGGTGGGAGCGCGGAGGAATGGA
 CATAAACCCGGGTGTGAAAGCCAGGGAATGAAGCCCGCGGAGCGGGGAATCTGGGATTCAGGG
 CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAGAGGGCAATCGCTCCTTG
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
 GGTGGAGTCAATACTGTGAACCTTGAGACGCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAGGACCGAGAGAAGCTCTC
 GCATCCCGTGAACCTCAGACAGGAGCCACAGATTGGATGGGAGCCCCACACTCCCTCCTTAA
 AACACCACCTCTCATCACTAATCTCAGCCCTTGCCCTTGAATAAACCTTAGCTGCCCCACAAA
 AA
 AA

FIGURE 52

MKFGQFLACLLALCLGSGEAGFLQSGEESTGTNIGEALHGLGDALSEGVGKAIKGEAGGAAGSKVS
 EALGCGGTREAVCTGVQRVFGPGAADALGNRVGEAAHALGNTGHEIGRQAEDEVIRHGADAVRGSWQGV
 GHSGAWETSGGHIFGSGGLGGQCGNPGGLGTPVWHGYPGNSAGSFGMNPQAPWQGGNGGPPNF
 GTNTQCAVAQPGYGSVRASNQNEGCTNPPSGSGGGSSNSGGGSGSQSGSSGSGNGDNNNGSSSGGS
 SSGSSSGSSSGGSSGGSSSGSSGNSGGSGRDSGSESSWGSSTGSSSGNHGSGGGNGHKGCEKPGNE
 ARGSGESGIQGRFGQGVSSNMREISKEGNRLGGSGDNYRGGQGSWSGSGGDAVGGVNTVNSETPGM
 FNFDTFWKNFKSKLGFINWDAINKQDRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
 90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
 159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
 236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
 253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
 283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
 298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
 325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGCTGCTGCTGAGCCTGCCCTGG
CTGGGCTTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCTCTCTATAACAACCTGCGCGCGCTCCAGTGTCTCC
CACAGCCCCAAAACGGAACCTGGTTTTTGGGTCACCTGGGCTGATCAGCTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTGCGCCACCTATCCAGGGCTTTACGGTATGGCTGGTCCCAT
CATCCCTTCATCGTTTTATGCCACCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCAAGGATAATCTCTTCATCAGGTTCTTGAAGCCCTGGCTGGGAGAAGGGATAGTGTCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGAGGACTCAGTGCCCTTGAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGCGCTTCCAGGGGCTGCCGCTGGTGCATGACTTACAGAC
GCTGTCATCCGGAGCGCGCTGCACCTTCCCACTCAGGGTATGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTCTTCTGCTGAGCAAGGATGAAGATGGGAAGS
CATTTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGAGGGCCATGACACCAG
GCCAGTGGCTCTCCTGGTCTGTGTACAACTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGGCC
AGCTGCCCTCTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC
TCCCCGATGTGACCCAGGACATGTTCTCCAGATGGCCGAGTCAATCCCAAAGGATTACCTG
CCTCATCGATATTATAGGGTCCATCACAAACCAACTGTGTGGCCGGATCCTGAGGTCACGACC
CCTTCGGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACTCTGGCTTTTATCTCTTCTCCGCA
GGGCCAGGAACATGCATCGGCGAGCGGTTGCGCATGCGCGAGATGAAGTGGTCTCGGCTTGAT
GCTGTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGCGGGCTTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGACAGTCACTTTCTGAC
CCATCCACCTGTTTTTTGACAGATTGTCATGAATAAACCGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGLRPFVAMSPFWLLLLLVGSWLLARILAWTYAFYNNCRRLQCFPOPPKRNWFVGHLG
 LITPTEGLKDSQMSATYSQGFTVWLGP IIPFIVLCHPDTIRSI TNASAAIAPKDNLFIRFLKP
 WLGE GILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQH LASEGSSRLDMFEHI
 SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFRHAC
 RLVDHFTDAVIRERRRTLPTQGIDDFPKDKAKSKTLD FIDVLLSKDEDGKALSDEDIRAEADTF
 MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
 LHP PAFP TISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
 LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
 VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGTGTGTTTGGCAGTTACCCCTGTGTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACCTATCTTCTGCATTTTGAGTCTTTGGCATTGACGTGG
TACAGCCTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTTTTGCCGTGTCTTGC
ATAATTTCATGGCCAGTTTTATGAAGCTTTGSAAGGCATATGGACAGAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGACAGCAATGTGTGCTT
GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAACTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCACTGGAGCACTCCAGCAGGCCGCAACATGCTCTGTCTGTGCTCTG
TAGCTGGCCGGTTCATCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTGGAGGGGCTCCCTGCCAGGCTGAAGTCT
ATTTCAGAGCTCAGTGTCTTCATCCCTCCAGGATTTCCAGCTACCGCCAGTGGAGAGCAAAAAATTGACAAAGCT
GGAGATAGAGCTTAGTGGGCACTTAGAGTTTGAAGAAATTTGTCAATTATCTCCAGATCATGAGAGAGAGCTGAGG
CTGGTGTTTAAGATTTTGGCAAAAGATATGATGGAGCTTAGTGGGCAAGGATCATCGAGTCCCTCGGGAGCTTG
GGAGTCAAGATATCTGAACAGCAGGCGAGAAAAAATTTCTCAAGACATGGATATAAAGCGGAGAGTGAACATGAGCTG
AACGAGTGGAGAGACTACCACTCTCCACCCGCTGAAAAACATCCCGAGATCATCTCTACTGGAAGCATTCACAG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGGAGCAGACGGGATGTGTGTGGAG
CACCTGTGGCAGAGTGGGGCAGGGGCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGGCTCATG
CAGGTCCATGCTCCCGCAGCAACCAATGGGCTGTTGGTGGCTTCACTCAGATGATTCGAGAGAGGAGGGGCCAGG
TCAGTCTGGCGGGGGAATGGCATCAACGTCTCAAAATTGCCCGGAATCAGCCATCAAAATTCATGGCTATGAGCAG
ATCAGGCGCTTTGTGGTGTGACCAAGAGACTCTGAGGATTCAGCAGAGGCTGTGGCAGGCTGTGGCAGGCTCTTGGCAGGGGCC
ATCGCCCAAGCAGCAGCTATCCCAATGGAGTCTCTGAGAGCCGGATGGGCTGACAAAGAGCTATGTCCCAACATGTG
ATGCTGAGCTGGCCAGAGAGATCTGGCAGAGAGGGGGTGGCGGCTTTGACAAAGAGCTATGTCCCAACATGTG
GGCATCATCCCTATGCCGAGCTGACCTTGCAGCTACGAGAGCCTCAAGAATGCTGGCTGCAGCAGATGTGAGTGTG
AACAGCGCGGAGCCCGGCGTGTGTGTCTGCTGGCTGTGGCAGCATGTCCAGTACTGTGGCAGCTGGCAGCTAC
CCCTGGGCTAGTCAGGAGCCGGATCAGGCGCAAGCTCTATTGAGGGGCTCCGAGAGTGACATGAGCAGCTC
TTCAAAACATATCTCGGCAAGGAGGGGCTTGGGCTGTACAGGGGCTGGCCCCCACTTCATGAAGTCAATCCCA
CTGTGAGAGCATCAGTACGTGGTCTACAGAGAACCTGAAGATCACCTGGGCTGAGTGGCTGGGCTGTGGAATGTGCCAACAT
CGCCCGGAGTGGAGCTGCTGATCTTGGGCGCAGGCTGGGCTGTGAGGAGGCTGGGAGAGCTGGCAGGCTCAGGGGCTT
AAGCTGTCTCGAGCCAGATGTTGAAAGCTTAGAGCCAGGAGGCTGGGAGAGCTGGGAGAGCTGGCAGGCTCAGGAG
GTCTGTGCTGACCCAGCAGACCTCCTGTGGTCCAGCAGAACACAGGCTCTTGAAGCTCAGGCTCAGGAG
CTCGGGCTCAGATGTGAAGCAGCAGCATTTCTGCACTGCTGCCAATATGAGCTTGGAGCTGGAGGCTCAGGAG
TAGCTTCTCCATTTCACTTCCAGCTGCAGCAGCTGTTGGCAGCGGCTTGGCTCTGGTCTGGCTGCTCTCTCTGTG
CTCTTGTGGCTGCTGTCTGTGAGTAAAGTGGGAGAGGCTCAGCAGCAGCTCCACCCCTGCTCAATCCC
ATAATCCATGATGAAGGTGAGTCACTGGGCTCCAGGCTGACTTCCCACTCAGCAGTGAAGCCCACTTGGC
TGTGAAGAGAGGAGAAAGATCTGGCTTGTGCTACTGCACTGAGCCCTGCTGATGGCTGGGCTCTCGGGATG
CTTGGAGATGAGGGCTCGGGCTGGCTGGCTGGCTGCACAGAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCTGCAAGCTCTCAGATGGGCGCAGCTCAAGCAAACTCACTGTCCCACTGTGGCATGAGGCGAGTGGAGCA
CCATGTTGAGGGGGAAGGCGAGAGCTTGTGTCTTGGGAGGAGAGAGAGGTGTGGAGGCTTAATATAGG
ACTGTTGGAAAGGGTTTTGTCAGAGAGGCAAGCCGAGCAAAATGAGGAGCTTGTGCTTCCAGAGAGAGGAGG
GAGCAGGAGGCTGGCTGACTGCTCAGAGTCTGTTCTGAGCCCTGGGGTCTGTGCCAAGCCAGCAGGCGCAGC
GGGACAGCCCAATCCATTTGTGCTGCTGCTGGAACCTATTATTTTGTATTATTTGAACAGAGTTATGTCT
AATATTATTATGATTTGTTAATTAAAGCTTGTCAATTTCAAGTTCAITTTTATTCATATTTATGTTCAATGGT
GATTTGACTTTCCCAAGCCCGCCAGTGGATGGAGAGGAGGAGAGGAGGGGCTTGGGCGCTCGAGTCACATCT
GTCCAGAGAAATCTTTGGAGCTGGAGGCAAGAAAGCGCCAGAGGCGCAGCAGCAGCTGGCTCTTCTTTGGCAG
TGTGGAGAGGCTTGGCCAGGCTTAGATTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTT
AACCTTGAAGTGGATCAATTTCTCTGGCTGGAGGTTTCAATTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTT
TGAAGAAATCTCAATAGATGAAAGTCAATGCAAAATTTGTTATATATGAACATATAACTGAGCTGTCAAAAG
CAATTAAGAAAGATTTGGAGCTTGAAGTTGTCTATTAAGCAGCTTCAATTAAGTTGTTTCAAGCTGAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIQSFSTYRQWKQKIVQAGDKDLG
QLDFEEFVHLYQDHEKKRLRVFKILDKNKNDGRIDAQEIQSLRDLGVKISEQQAEEKILKSMKNG
TMTIDWNEWDRDYHLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWRHLVAGGG
AGAVSRTCTAPLDRKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
FMAYEQIKRLVGSQDETLEIRHERLVAGSLAGNIAQSSIYPMEVLKTRMALRKTQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNANLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVTRMRMQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVISISVY
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
 TTCCCTGGGGCAGATCCTCTTCTGGAGCATAAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
 CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCAGACTCACTACTGTGCGCTCAGCT
 GGGACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTGAACCTGACATCAAACTTTCTGATAT
 CGTGATACAATGGCTGAAGGAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAGATG
 AGCTGTGCGGAGCAGGATGAAATGTTCAGAGGCCGACAGCAGTGTTTGCTGATCAAGTGATAGTT
 GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
 CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
 AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCGATGGTTCCCC
 CAGCCACAGTGGTCTGGGCATCCCAAGTTGACCAAGGAGCCAACCTTCTCGGAAGTCTCCAATAC
 CAGCTTTGAGCTGAACCTCTGAGAATGTGACCATGAAGGTTGTGCTGTGCTCTACAATGTTACGA
 TCAACAACACATACTCCTGTATGATTGAAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
 ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAACTCAAGGCTTCTCTGTGTGT
 CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
 GTGCCCTTGCCCAAAAAAGCATGCAAAGTCAATTGTTACAACAGGGATCTACAGACTATTTCAC
 CACCAGATATGACCTAGTTTTATATTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
 AGCAAAACAGAGCAAGAAACAAAAAGAGCCAAAGCAGAGGCTCCAAATATGAACAAGATAAAT
 CTATCTTCAAAGACATATTAGAAGTTGGGAAATAAATCATGTGAAGTACAGCAAGGTGTTAAGA
 GTGATAAGTAAAAATGCACGTGGAGACAAGTGCATCCCAGATCTCAGGGACCTCCCCCTGCCTGT
 CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTATGTTATATGTGC
 TGTAAATGTGCTCTGAGGAAGCCCCGGAAGTCTATCCCAACATATCCACATCTTATATTTCCAC
 AAATTAAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCACTCAGGGGGG
 GCTGCATTTTAGTAATGGGTCAAATGATTCACCTTTTATGATGCTTCCAAAGGTGCCTTGCGCTC
 TCTTCCCACCTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
 CGGGGACACCGATTTTATAAATAAATGAGCACCTTCTTTTAAACAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLLKNVQLTDAGTYKC
YIITSKGRGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNNTYSCMIENDIAKATGDIKVTESEIKRSHLQLLNKASL
CVSSFFAISWALLPLSPYMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGCAGGGCTGCCAGGAGCTGCAGGACAAG
 CACCAGSAGCCCCCTCCGGGTAGCTACTACCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
 GGGCTACCCCCCTGGTGGTGGTTATGGGGGTCTGCCCTGGAGGGCCTTATGGACCAOCAGCTGG
 TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAAGCTCCAGGAGGACCATATG
 GCGGTGCAGCTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCTACGGTGCCCGAGCAGCCT
 GGGCTTTATGGACAGGGTGGCGCCCCCTCCAATGTGGATCCTGAGGCCCTACTCTGGTTCAGTC
 GGTGGACTCAGATCAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
 ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
 GGCCGCATCGATGCTACGGCTTCTCAGCCCTGTGGAAATTCACTCAGCAGTGGAAGAACCCTTT
 CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
 AAATGGGCTACAACCTGAGCCCCCAGTTCAACCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
 GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTCACCCAGCTGCAGGTGCTGACAGA
 GGCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
 CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
 TCCTGCCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTTCTTTCTGTCCCTCTAGAAGAAC
 ATTCCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGGTCTCTGCATCATAGCCACCA
 AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCCTGATGGAGGAGAGGATAGAAGTTGA
 ATGTCTGATGGCCATGAGCAGTTGAGTGGCAGAGCTGGCACCAGGAGCAGGTCTTGTAAATGG
 AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAAGTTCATCGGCCCTGTTACC
 GTTAGTACCTGTGTTCCCTCACCAGGCCATCTGTGTCAAACGAGCCATTCTCTCCAAAGTGGAAT
 CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTTGGATTCTGCCACACCCATANAAT
 CCTTGTGTGTTAACTTCTAGCTGCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
 CTTTGCCAGGCCTTCTGCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTGGCT
 TCAGTCTCCAGGAGCAGTGGTACCTCTCCCTGCCAATACTTTTTTAATTGCATTTTTTTTC
 ATTTGGGGCCAAAAGTCCAGTGAAATTGAAGCTTCAATAAAGGATGAAAACCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYFPGPFNSGGQYGSGLPFGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTFGGPGYGGAAPGGPYGQPPSSYGAQQPGLYQGGAPPNVDPPEAYSWFQSVDSDH
SGYISMKEKQALVNCINWSSFNDETCLMINMFDKTKSGRIDVYGFSA LWKFTIQWKNLFQQYDR
DRSGSISYTELQQALSQMGNLSPQFTQLLVSRYPNSANPAMQLDRFTQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDEFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCGCTGCTCCTGGTCTGTCTTCATC
 TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGGGAGGAGAAAGTTTCCCAAACATTCGGGACC
 AACTTGCTCAGCTCGGACACCTTCCTCCACTGGCCCCCTTAACCTCTGAACATCOGCAGCCCCG
 TCTGGACCCCTAGGTCTAATGACTTGGCAAGGGTTCCCTCTGAAGCTCAGCGTGCTCCATCAGATG
 GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
 GATTCTTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
 AGCGCTGCTGAAGAACTCTCTTACCTCTCCAGTCTGCGGCCCTCGCTCCGGGAGTGGCCCTT
 TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
 TCGGAGTCCAGACGACTGCCCCGTCTTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
 CCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCACTG
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCATGCCACACCTGAGGGA
 ATCTGGGGTATCAATAATCAACCCCCAGGTACAGCTGGGGAAATATTAAATCGGTATCCAGGAGG
 CAGCTGGGGAAATATTAAATCGGTATCCAGGAGCAGCTGGGGAAATATTAAATCGGTATCCAGGAG
 GCAGCTGGGGAAATATTCACTATACCCAGGTATCAATAACCATTTCTCCTGGAGTCTCCGC
 CCTCCTGGCTCTTCTTGGAAATCCAGCTGGCTTCCCTAATCCCTCAAGCCCTAGGTTGCACTG
 GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCTGCTCCGCCCTCTGCTG
 TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTTCAGTGCTCCCC
 TGCTCATCTCAATAAAATAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 64

MQGRVAGSCAPLGILLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQFSSTGFSNSEHPQPAI.
DPRNSDLARVPLKLSVPPSDGFFPAGGSAVQRWPPSWGLPAMDSWPPDPWQMAAAAEADRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRNSNLGAGGKILSQRP
PWSLIHRVLPDHPWGTNLNPSVSWGGGGPGTGWGTRMPHPPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNI PAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCAGGAGCGCAGTGGCCACTATGGGGTC
 TGGGCTGCCCCCTTGTCCTCTCTTGACCTCCTTGGCAGCTCACATGGAAACAGGCGGGGTATGA
 CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
 TTGCTTGAAGGCTCTGCCTCCTCCTCCATCTCCCTTCAGGACCCAGCGTCACCTCCACCATGC
 AAGATCTCAACACCATGTTGCTGCAACACATGACAGCCATTGAAGCCTGTCTCTTCTGGCCC
 GGGCTTTTGGCCGGGGATGCAGGAGGCAGGCCCGACCTGTCTTTCAGCAGGCCCCACCTC
 CTGAGTGGCAATAAATAAAATTCGGTATGCTG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
219

FIGURE 66

MGSGPLVLVLLTLLGSSHGTGPGMTLQKLKESFLTNS\$YESSFLELEKLCLLLHLPSGTSVTL
HEARSQHVVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGTTGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCCGTGCGCAGGTGCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCTCGGGCTGCCAACCCGCCACCCAGGCCATGGCGAACCCTGGGCTGGGGCTGCTTCTGGCG
CTGGGCTGCCCTTCTGCTGGCCCGCTGGGGCCGAGCTGGGGCRAATACAGACCACTTCTGC
AAATGAGAAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCCTTCTCTGGCTGTGGGGCTG
GCACTGTTGGTGCGAAGCTTCGGGAGAAGCGGCGAGACGGAGGGCACTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGTTGCAGGGCT
GCCTGCCCATCTAGGTCCCCCTCTCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGTACTTCRA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLRWGRWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEQFSHAAEARAPQDSKETVQCCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATACTAGAGAGGAACAAATGGGTTATTCCAGAGGTTTGTTCCTCTTAGTCTGTGCTGCTGCACGAG
 TCAATACTCTCTCTATTAAGTGTAATAATATGGCTTTGAGGATATTTCATTGTATAGATCTCAGTGTGCCAGGA
 GATGAAGAAATAATTGACAAATAAGGATATGGTGGCTACAGCTCTGAGTACTGTTTGGAGCCACAGAAAGA
 TTTTCTTTCAAAATGTATCTATAATTAATTCGCTGAGAAATGGAGAGAAATCTCAGTACAAAGGCCAAACATGAA
 AACCATAAACATGCTGATGTTATAGTTGGACCACTACACTCCAGTAGAGATGAACATACACCAAGCAGTTACCA
 GAATGTGGAGAGAAAGGCCAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAATGAATATGGACCA
 CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACTCCGCTGGGGAGTGTGATGAGTACAAATGAAGATCAGCCTTTC
 TACCTGCTGAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGAGGTATCTCTGTTAGAAATAGAGTTTATAAGTGT
 CAGAGGAGCAGCTGTCTAGTAGAGCATGAGAAATGATTTACACAAAACTGATGAAAAAGATTGTCAATTCCTT
 CTGATAAATGATCAACAGAAAGAAATGATATGTTATGCAAGTATGATTCCTGTTTGAATTTGTAAAGAA
 AAAACCAATGATCAAGAGCTCAAGCCTACAAACATAAAGTCCAATTTAGAGTACATGGAGAGTATGAGCAAT
 TCTGAGGATTTAAAAACCATACCCATGGTGACACCACTCTCCACCTGTCTTCATTGCTGAAGATCAGTCAA
 AGAATTTGTGCTTAGTTCTTGATAAGCTCGAAGCATGGGGGTAAAGACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTCTGTCGAGACTGTTGAAATGGATCTGGTGGGATGGTTCACTTGAATGACTGCCCATTTGTA
 AATAGCTAATCCAAATAAAGCAGTGTGATGAAGAAACACACTCATGGCAGGATTAACATATCTCTCTGGAGGA
 ACTTCATCTGCTCTGGAAATAAATATGCATTTCAAGTGTATGGAGAGTACATTCACCACTGATGATCGGAGATA
 CTGCTGCTGACTGTGGGGAGGATAACACTGCAAGTTCTTTATGATGAATGAACAAATGGGGCAATTTGTCAT
 TTTATCTGTTGGAGAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAGAGTAACAGAGGAGAGTCATTTATGTT
 TCAGATGAAGCTCGAAGCAATGGCTCMTGATGCTTTTGGGGCTCTTACATCAGGAATACTGATCTCTCCAGAGAG
 TCCTTTCAGCTCGAAAGTAAGGATTAACTGATGAATAGTAATGCTGATGAACGACTGCTCATATATGATAGTACA
 GTGGGAAGGACACGTTCTTCTCATCATCTGGAACAGCTGCTGCTCCAGTATTTCTCTCTGGGATCCCAAGTGAACA
 ATATGAGAAATTTACAGTGGATGCACTTCCAAATGGCTATCTCAGTATTTCCAGGAATCGAAGGTGGCCT
 TGGGATGATATCTTCACCAAGAGCAACCAAGCAATTAATCATACAGTACTCTCGAGAGCAATTTGGCTCT
 GTGCTCCCATCCAGCTGAATCTCAATGAAATAGGAGGTAAACAGTTTCCCAAGCCATGATTTTACGAGAA
 ATTCTACAGGATATGATCCTGTTCTTGGGCCAATGTGACTGTTTCATGAATCAGAGATGGACATACAGAGTT
 TTGGAATCTTTGGAATATGGTGAGGCGCTGATTCCTTCAGAAATGATGAGTCTACTCCAGGTATTTACAGCATAT
 ACAGAAATGGCAGATATAGCTTAAAGTTCCGGCTCATGGAGAGCAACACTGCAAGCTAAATTTACGGCTTCA
 CTGAATAGAGCCGCTACATACAGGCTGGGTAGTGACAGGGGAATTTGAAGCAAAACCGCCAGACCTGAAATGAT
 CAGGATCTCAGACCACTCTGGAGGATTTCAAGCCAGCAGCAGTCAATGCTGATGAGTATGAGATTTCCAGC
 CTCTGCTTCTGCTGACATACCCAGCAATCAATACAGACCTTGATGGCAGGCTCATGAGATTAAGATTATCTT
 ACATGGACAACCCAGGAGATAATTTGATGTTGGAAAGTTCAAGCTTATATCATGAATTAAGTCCAGATATCTT
 GATCTAGAGACAGTTTGTATGATGCTCTTCAAGTAAATACTACTGATCTGTCCAAAGAGGCCCCAATCCAGAA
 AGCTTTGATTAACCCAGAAATATCTCAGAAAGAAATGCAACCCACATATTTATTGCCATTAAGATATGATATAA
 AGCAATTTGACATCAAAAGATCCACATTTGCACAAGTAAGTTTATTCTCCCTCAAGCAATCTGATGATCATGAT
 CTTCACCTACTCTCTACTCTCTACTCTCTGATTAAGTCAATTTCTGGAGTAAATTTCTACGCTGGTATGTC
 TCTGATTTGGTCTGTGTAAATTTGAATTTTAAATACCACTTTGACCTTAACGAGAGAAAAATCTCT
 AAGTAGACTAGAGAGGTTTAAACAAACATGTAAGTAAGAGATTTCTGAATCTTAAATTCATCCAT
 GTGTGATCAATAACTATAAAATTAATTTAAGATCTCGGAAGAGATACCTTGATTAATAAAACACTCATGGATA
 TGTAAAACTCTCAGATTAATAATTAATAGTTTCAATTTATTTGTTATTTTATTGTAAGAAATGATGATGACAAAG
 ATCTTTTTCATCTGATACCTGGTGTATATTTTGTGACACAGTTTCTGAAATGATATTTCAAAATTCATCAA
 GAAATTAATCATCTATCTGATAGTCAAAATACAGATTAAGGAGAGCAATAAACACATTTTGAIAAAAAA
 AA

FIGURE 70

MGLFRGFVFLVLVCLLHQSNSTFIKLNNGNFEDIVIVIDPSVPEDEKIEQIEDMVTTASTYLFE
ATEKRFFFKNVSLIPENWKENPQYKRPHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEQPFFRAKSKKIEATRCISAGISGRN
RVYKCGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVINSSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSGMGKDLNR
MNQAAKHFLLQTVENGSGNVGMVHFDSTATIVNKLIIQIKSSDERNTLMAGLPTYPLGGTSCSGIK
YAFQVIGELHSQLDGSEVLLLTGDGENTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFVVSDEAQNGLIDAFGALTSGNTDLSQKSLQLESKGLTLMSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLITITVTSR
AANSSVPPITVNAKMNKDVNSFPPSPMIVYAEILQGYVVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNQGVYSRYFTAYTENGYSYSLKVRAGGANTARLKLRLPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLDEFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILWTAPGDN
FDVGKQRYIIRISASILDLRDSFDDALQVNTDLSPEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSINQAVTLFIPQANPDDIDPTPTPTPTPTPDKSHNSGVNISILVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCTTTAAGTGGAAACCTGGGAGTAGAGTACTGACAGCAAGACCGGGAAGACCATACGTCCTCCGGGCGAGGGGTGA
CAACAGCTGTACTCTTTTGATCTCGTGTGGGCTGCTTCTATTCAAGGAAGACGCCAAGGTAAATTTTGACGA
GAGGAGCAATGATGTAGCAACCTCTTAACCTTCCCTCTTGAACCCCAAGTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACACGACNAGCGGCTCTTCCGCTTAACCTTGTGGTTGGAGGAGAAACCTTTGTGGGCTGGGTTCTCTTAGCA
GTGCTGAGAGTGAAGTCTTGGCTGGGGGTGACAGAGAAAGGAAAGGTCCTCTGCTGTGGCTGCATCAGGAA
GGCTGTATGGGAATGAAGGTGAAAATCTGGAGATTTCACTTCAGTCAITGCTCTGCTGCAAGATCACTCTTTAAA
AGTAGAGAGCTGCTCTGTGTGGTGGTTAACTCAAGAGGCAAGACTGTTCTAGAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCAACAGCATGCTCTCTGTGGTGTACGCCAGGAGAGCCCTTCCCTGGGGGCCCGGCTTGAGGGATGCC
ACGGGTTCTGAGCAGATGGCTATCTCTGAATGATGGTTCCGCGGGGCTGCTGGTGGGATTTCCGGGGTGGG
GTTTGTGCTGGTGCCTCTGCTGTGCTATCTCTGCTCTGTACATTTGGGCTGCACCCCAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGCGCAACAGCCCCACGGGGAAGGGGTACCAAGCCGCTCCTTCAGAGTGGGAGGACGACGAC
CGCACTACTGTGACGAGCTGAGGCGGAGATGCGACAGCTCAAGGAGGAGCTCGAGGAGAGGAGTGAGCAGCTCAGG
AATGGCGAGTACCAAGCCAGCATGCTGCTGGCCTGGGCTGGAAGAGGCCGCCGAGAGAAAACCCAGGCCGACCTC
CTGGGCTCTTGCACCTCGCAGGTGGCAAGGCAAGGTGAATGCTGGGCTCAAGCTGGCCACAGATATGACGAGTG
CCTTGTGATGACTTTACTCTACGAGAGGTGTACAGCTGAGAGACTGGCTTACCCGCCACCCGAGGAGAGGCTGTG
AGGAGGAGCAAGGGGATGAGTGTGGGAAGCATTAATCAGCCTTGGAGACTTGACATCTCTCGAGACACAGC
CCCAATCACCGCTCCTTACAAGCGCTCTGATTTTCAAGAAGGATCTACGACACAGAAAGGACAAAGGACATTTGAT
GAGCTCACCTTCAAAGGGGAACCAAAACAGGAATCAAAGGCTCATCTTATTCGACATTCAGCCGCATCATGAAA
GTGAAAATGAAAAGCTCAACATGGCCAAACGCTTATCATGTTATCTGTGCTCTAGCAAAAAGGGTGACAAAGTCT
CGGAGTTCTAGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGTCCATCTCACTGTTTGTATCTTTGGG
AAGAAGAAATAAATGAAGTCAAGAATACTTGAANAACACTTCCAAGGTGCCAACTTCAGGAACCTTACCTTCATC
CACTGAATGAGGAATTTTCTCGGGAAAGGAGACTTGATGTGGAGGCCGCTTGGAGAGGAAGCAAGCTCTTCTC
TTTTCTGTGATGGAGTCACTTACTTTCACTCTGATGCTCATAGCCTGTAGCTGATACACAGCAAGGAGAG
GTATTTTATCCAGTCTTTTCACTGATGATACATCTGGCATAATACAGGCCACCATGATGCAAGTCCCTCCCTTGGAA
CAGCAGCTGTGCATAAAGAGGAACCTGGATTTTGGAGAGACTTTGGATTTGGGATGAGCTGACATGATGATGACAG
TTCACTCAATATAGTGGGTTGATCTGGACATCAAGGCTTGGGGCGGAGAGATGTGCACCTTTATCCGAGTATCTC
CACAGCAACCTCATAGTGTGTAAGGAGCCCTGTGCGAGGACTCTTCCACCTTGGCATGAGAAAGCGCTGCATGGACAG
GTAGCCCTCCGACAGTACAGATGTGCATGAGCTCCAAGGCCATGAACGAGGCTATCCACGGCCAGCTGGGATGCTG
GTGTTTGAGGCAAGGATGAGAGCTCACTCTTGGAAACAGAAACAGAGCAGTATGCAAAAACAGCACTCCCA
GAGAGATGTGGGAGCACTTTTCTTCTTTGCAATCTAGAACTGGGCTCACAGAGAAAGACTTCCCTAAA
GGACGACAAAGAATGAGCTATGGGTGAGAGATGAGAAGCCCTCGATTTCTCTGTGGGCTTTTACAACAGA
AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCACCTGTGAAGTGTGTGACAAAGGCAGATGCTTTGTG
AGATTATAAGCTAATGGTGTGGAGGTTTGAATGGTGTTTACAATAGCATGAGACCTGTTGTTTGTGTGCTCATGTGA
AATATTCACTGATTAGAGCAGTTTGTGAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCTCATATGAATGA
GCTATTCAAGAGGCTCTAGTTCTAGGAATGTGAAATATCAGAAAGCAGGAGAGGAGATAGGCTTATATGATACT
AGTGATGATTAAGTAAATAAATAGGACGAGAAAGAAAGCAATTAATGATGCTCTATTTTTCCCAAGAT
TAACCAAAATATCTGCTATCTTTTGGTGTCTTTTAACTGTCTGCTCTTTTCTTTTCTTTTAAATAGCAT
TTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATACCACCTTGCNAGCCTTACAAGAGGACAGTGGGCTAC
ATTTTTTATTTTAAAGAAGTACTTTGGAGTCATTATGAGAATCTTCACTCAAAGCATCAATTTATGCTCAT
CCAGGAGCATGCCAAATGCTGATTTCTGTGAGGCTCAATGATGAGGCAATGAGACATAGGGAGGAATGTTGTAT
AATACAGCTGCAGATACTTTCTGGAAGATTTTGAAGAGGAGCACTGAACACTGGGAGGAAGAAATGAC
ACTTTCTGCTTTACAGAAAGGAACTCAATTCAGACTGGTATGCTGATGATGCTTAAAGTCAGAAACCACTTTT
CTCTGAGAGTAAAGGACCTTTCTTACTGTTTAAATAAACCAAGTGGAGCTGTGAGACCAACATCTTTT
AAAACAGGGTGCTCTCTCGCTCTGCTTCCATAGAAGAAATGGAGAAATATATATATATATATATTTGT
GAAGATCATTCATCTGCCAGATCTAGTGGGATGGAAGTTTGTGTCATGTTATCAACCAGGCGGAGGTGGAG
TAACGATTAATTTTAAATTAAGCACTTACTCATCAACCAAGATGCTGTGAAAATGCAATTTATACCAATTT
CAAACTATTTTAAAAATAAATACAGTTAATAGAGTGGTCTTCACTCATGTGAATAATTAAGCAGCACCAG
ATGCATGAGCTAATATCTCTTTGAGTCTTGTCTCTGTTGCTCAGATAACTCATTTGTTAAAAAGCTTCAAGAC
ATTCAGGCTTGGTGTGTGTTAAAAATGCAATGATTTGATTTGACTGGTATTTATGAATTTAATTAACACAGG
CCATGATGGAGTGGATTGCGACAGCTAATAAATATGATTTGGTGGATGA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQALPRANSP TGKEGYQAVLQEW E
EQHRNVVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAVGKLATEYAAVFFDSFTLQKVYQLETGLTRHPEEKFVRKDKRDELVEAIESALETLNPA
ENSNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPI MKVKNEKLMAN
TLINIVIVPLAKRVDFKRFMQNPREMCIEQDGRVHLTVVYFGEKEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDVIYFTSEFLNTRCLNTQPGKKVFYFVLFSQY
NPGIYGHHDVAFPLEQQLVIKKE TGFWRDFGFGMTQCYRSDFINIGGFDLDIKGWGGEDVHLYR
KYLHSNLIVRTEFVRGLFHLWHEKRCMDELTP EQYKMCMQSKAMNEASHGQLGMLVFRHETEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

[illegible]

1

FIGURE 74

MLFSALLLEVINILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGSEHQINSEAT
FAELHIVHYDSDSYDSLSSAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMFASFIAQGSSYTTGEMLSLGVGILVGCCLLLAVYFTIARKIRKKRLNRRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCTGCGCCGCTGCTGCTGTTGCTCTCGCGGCGCTTGGGAGCGGCAGTTCCTGTGTCT
 TCTGGTGGTTTGCTTAAACCTGCAACATCACCTTCTTATCCATCAACATGAAGAATGCTCTACA
 ATGGACTCCACAGAGGGCTCTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
 ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTGTCTGACAGCTCC
 AGAGAAGTGGAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
 AGTATAACGTGTCTGTGTTGAATACTAAATCAACAGAACGTGGTCCAGTGTGTGACCAACCAC
 ACCTGCTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCTCTGCTCCC
 AGGGCCCCCTCGCGTGTCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
 CAGAGTTCAGGCTAAATCATCTTCTGGTATGTTTTGCCATATCTATTACCGTGTTCCTTTTT
 TCTGTGATGGCTATTCCATCTACCGATATATCCACGTTGGCAAGAGAAACACCAGCAAAATTT
 GATTTTGATTATGGAATGAATTTGACAAAAGATTCTTTGCTGCTGAAAAATCGTGATTA
 ACTTTATCACCTCAATATCTCGGATGATTCTAAAAATTTTCATCAGGATATGAGTTTACTGGGA
 AAAAGCAGTGATGTATCCAGCCTTAATGATCTCAGCCCGCAGGGAACCTGAGGCCCTCAGGA
 GGAAGAGGAGGTGAACATTTAGGGTATGCTTCGCATTTGATGGAATTTTTTGTGACTCTGAAG
 AAAACACGGAAGTACTTCTCTCACCCAGCAGAGTCCCTCAGCAGAACATACCCCGATAAA
 ACAGTCATTGAATATGAATATGATGTGAGAACCAGTGCATTTTGCGGGGCTGAAGAGCAGGA
 GCTCAGTTTCAGGAGGAGGTGTCACACAGGAACATTATTGGAGTCGAGGCAGCGTGGCAG
 TCTTTGGCCCGCAACGTTACAGTACTATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
 CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGGCCATCGACGACCTGGTGCAGTGGGA
 TCCCCAACTGGCAGGCTGTGATTCTTCTGCTGTCAGCTTCGACCAGGATTACAGGGCTGCG
 AGCCTTCTGAGGGGATGGGCTCGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGT
 CCAGACAGGCCACAGGAGAAATGAACCTATCTCATGCAATTCATGAGGAATGGGGTTATA
 TGTGCAGATGGAACTGATGCCAACACTTCCTTTTGCTTTTGTTCCTGTGCAACAAAGTGAG
 TCACCCCTTGATCCAGCCATAAGTACTGGGATGAAAGAAATTTTTCCAGTTTGTGAGTGT
 CTGTGAGAATTACTTATTCTTTTCTTACTTCTATAGCACGTGTGTGATTGGTTATGCAATGTA
 GGTCTCTTAACAATGAGTGGGCTCTGGAGTCCAGGGGCTGGCGGTGTTCTATGCAAGAA
 AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELALLTLCSSSQIGPPEVALTTDEKSISVVLTAPEKWKRPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAPQSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFPDKRFFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEDVRTTDICAGPEEQELSLQEEVSTQGTLLSQ
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPAFDRPPGENETYLQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGGCCAGGTCTGGCATCTGCAC TTGCTGCCCTCTGACAC
CTGGGAAGATGGCGGGCCCTGGGACCTTCACCC TTCTCTGTGGTTTGTGGCAGCCACCTTGATC
CAAGCCACCTCAGTCCCATGCGATTCTCATCTCGGCCCAAAGTCATCAAAGAAAGCTGAC
ACAGGAGCTGAAGGACCACACGCCACCAGCATCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCAACAGCTAACATCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGTAGTCAAGATCCCTCGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGAGGACTGAGGCCCAAGCCACCATCCGCATGGACACAGTGCAAGTGGC
CCCACCCGCTGGTCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAC TGTGTGA
TAAGCTCTCCTTCTGTGGTGAAGCCTTAGCTAAGCAGGTCAAGAAC TCTAGTGCCATCCCTGC
CCAATCTAGTGA AAAACAGCTGTGTCCCGTGATCGAGGCTTCTTCAATGGCATGTATGCAGAC
CTCTGCAGCTGGTGAAGGTGCCATTTCCCTCAGCATTGACCGCTCGGAGTTTGACCTTCTGTGA
TCCTGCCATCAAGGGTGACACCATTCAGCTCTACCTGGGGGCCAAGTTGTTGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCTGGACACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTATGCTCCTGTTGGACTCTGTGCTTCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGTGGCCCAACTGATCGTGCCTGGAAGTGTT
TCCCCTCAGTGAAGCCCTCCGCCCTTTGTTCACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAGGTGAACCACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACCTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCGAACAGAAATGGCAAATTAAGATCTGGGGTCCAGTGTCATTGGTGAAGG
CCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGAAAACCAAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
GTCCAGCTGGGAGTATGGGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
CCTGTGAAAA

FIGURE 78

MAGFWFTTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDNATSI LQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTI VE
FHMTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYP AIKGDTIQLYLGA KLLDSQGKVT
KNFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVA AVLSPEEFMVLLDSVL PESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTP EFFIDQGHAKVAQLIVLEVFPSS EALRPLFTLGIEASSEAQFYT
KGDQLIINLNNISSDRIQLMNSGIWFPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLT PASLNKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGACTCTGAATGAGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCGCGATGGCCTCT
CTTGGCCTCCAACCTGTGGGCTACATCTTAGGCCCTTTGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAAAAGTTCTTATGTCGGTGCCAGCATTTGTACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCTTCTGGGCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCTCCCTGGCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGATGGCGGTAGCAGGTGGAGTCTTTTCATCCTTGGAGGCCCTCCTGGGATTC
ATTCCTGTTGCTTGAATCTTCATGGGATCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGATCATCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAACCTACTACGATGCCATC
CAAGCCCAACCTCTTGCCACAAGAGCTCTCCAAGGCCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCTACAGCCTGACAGGATATGTGTAAGAAACAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCGAGGGCCACAGGTGAGGGACACTACCCTGGATCGT
GTCAGAAGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAGGACAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATGGCAAGGATGCTCGCCATGCCAGCCITTCGTTTTCC
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCTCAACTTGAACCCCATTCCTTAAAGCA
GGACTCAGAGGATCCCTTTGCCCCCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCACTGACTGACCCCTCTGTGATCAAGACCCCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAGCAGTGGCTTTTGGGCATTGCTCTAACTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTATGACTCCACAGTGCCA
GACTAATTTGTGATGAAGTGAATAAAAACCATCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGLTVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAIISSLACIIISVVGMRICTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPIVANNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRNSNY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

FMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGTGGACCTTCCTTGTCTCTCCATCTCTCCCTCCTTTT
 CCCGCGTTCCTTTCCACCTTTTCTCTTCTCCACGCTTAGACCTCCCTTCTTGCCCTCCTTTTCT
 GCCCACCGCTGCTTCTTGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGATCTGTGGCCCTGTGCCTCCGTGTCTTTTGTCTCCCTTCTCCGACTCCGCTCCCGG
 ACCAGCGGCTGACCTGGGAAAGGATGTTCCGAGGTGAGGTCTCTCTCTGTGTGGGA
 CTCGCGCTGCTGTGGTTCCCTTGGACTCCACGCTCGAGCCGCGCCAGACATGTTCTGCCTTTT
 CCATGGGAAGAGATACTCCCCGCGGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCTGATGT
 ACTGCTGCGCTGTACCTGCTCAGAGGGCGCCATGTGAGTTGTTACCGCTCCACTGTCCGCT
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGCCAAGTGTGGAACTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAGTCTGCCAGCACACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCCAACAGTGTGCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCTCACAACTGCCCGGAACAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCCTGCTGCCAAGCTGCAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGGTGAGACTCCTCAGGATCCATGTTCCAGTGTGCTGGGAGAAAG
 AGAGGCCCGGGCACCCAGCCCCACTGGCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
 CAGACCCAGGGAGCAGGAGCACAACCTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
 GTGTGCATGGCGGGAAGACTACTCCACGGGAGGTGTGCCACCCGGCCTTCGCTGCTTCGGC
 CCGTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTCAACCCGAGAAGTGGCTGGGAAGTGTGCAAGATTGTGCCAGAGG
 ACAAGCAGACCTGGCCACAGTGAGATCAGTTCTACCAAGTGTCCAAGGCACCGGGCGGGTC
 CTCGTCCACATCGGTATCCCCAAGCCAGACAACCTGCGTGCCTTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAACTTCTTCACTTGAATCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCAGAAAGAGGCACAGCACTTCCGACTGTCTGCTGGCCCCACGAAGGTCACT
 GGAACGCTCTTCTAGCCACAGCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAGTGACCAAG
 ACATAACAAGACCTTAAAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTATAAA
 TAAGAAGTTGCATTACCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSLLGLALLWFPLDSHARAPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPVHCPQPVTEPQQCCPKCEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSGCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGGTTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY
HGEVWHPAFRAFGLPCI LCTCEDGRQDCQRVTCPTETPCRHPKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCGGTACGCGAGAGCCTCTCC
 GTGGCTTCGCACTTGGACATTAGGCCAGTTCTCCTCTCTCTAATCCATCCGTCACTCTCTCTGTCA
 TCCGTTTCCATGCCGTGAGTCCATTACAGAACACATCCAGTCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCTCTAAGCTGGGATCAGGCCAGTGGCAGGTGTTGGGCCAGACAGCCTGTCCAGGCCCTTGGTGGGGGAG
 GACGCAAGCATCTCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTCCGCTTCTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGAGGACAGCCATTTATGAGATGCCACAGTATC
 AAGGCAGGACAAACTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTGAGGCTGGAAACATTACT
 GTGTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCACAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTGAGCACTGGGCTCAGTTCCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCACTCTCGGGCTGTTCCCGGGCCCAAGCGAAGTGAAAGGTCCACAGGACAGGATTGTGCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGGAGATCTCTCTGACCGTCCAGAGAA
 CGCCGGAGGCATATCTGTTCCATGCGGCAATGCTCATCTGAGCCGAGAGGTGGAATCCAGGCTACAGATAG
 GAGATACCTTTTTCGAGCCTATATGTTGGCACCTGGCTACCAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTGGCAATTGTGGACTGAAGATTTCTTCTCCAAATTCAGATGGAAATCCAGGCGGAATCTGACTCT
 GAGAAAGAACACGACAGCCAGAAATTGAGAGACGCCCGGAACACGAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAACTGTAAACCATAGAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAGAGATTTACAGGAGAGAGTGTGGTGTCTCAGAGTTTCCAGCAGGGAACATTA
 CTGGGAGGTGACCGAGGACACANTAAAGGTGGGGGTGGGAGTGTGCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGATGATGGAGAACATTTGTATTTTC
 ACATTAATCCCCGTTTATCAGCGTCTTCCCGAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA
 TGAGTCTGGGACCATCTCCTCTTCAACATAATGACCAAGTCCCTTATTATACCTGACATGTCGGTTTG
 AAGGCTTATTGAGSCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAATCTCCATAGTCTATCTCC
 CCAGTCACCCAGGARTCAGAGAAAGGCGCTTGGCAAGGGCCTCTGCARTCCAGAGACAGCAACAG
 TGAGTCTCTCCTCAGGCAACACAGCCCTTCTCCCGAGGGTGAATCTAGGATGAATCAGATCCGACAT
 TCTTCTTTAGGGATATTAAAGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCGCGCCAAAGTGGCTTCCA
 GATGAAGGGGAGTGGCCTGTCCACATGGGATCAGGTGTCTGCTGCTGCCCTGAGCTGGAGGGGAGAGAGG
 CTGACATTTACATTTAGTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAATACCACTCTCAGGTGAG
 AACGCTCAGGAATTCATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATATGCTTAG
 ATCTATTATGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MAIMLSLVLSLLKLGSGQWQVFGPKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAAGLYGCRISQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFFRPTAKWKGPGQDLSTDSTRNDRMDHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
PQWKIQAEILDWRKKGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQSFQAGKHWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLIPT
LNPRFISVFPRTFPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQSEKEASWQRASAIPESTNSSESSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCACAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGCAAGTAACTGCTGACGATGCAGAGTTCGGTGA
CGGTGCAGGAAGGCCTGTGTGCCATGTGCCCTGCTCCTCTCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGCCAAATACAGACCAGGATGCTCC
AGTGGCCACAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCTTGAGCATCAGAGATGCCAGAAGAAGTGATCGGGGAGA
TACTTCTTTGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCAGGACCCCTGGAGTCCGGCTGCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCGCTCCTCGGTGCTCACCCCTCATCCACA
GCCCCAGGACCATGGCACCAGCCTCACTGTGAGTGACTTCCCTGGGGCCAGCTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCCTACCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCAGAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGCTCACTCCAGAGGGCCAGTCTCT
GGGCTGTGCTGTGCAAGTTGATGCAAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCTTGACCTGTGCCCTCACAGCCTCAAAACCGGGGTGCTGGAGTGCTTGGGTGCAC
CTGAGGAGTGACGTGAATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAGCCACATCAGAGTGACTCAGGGGTGGTGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCCCTCTGCGTCACTTTCGTTGTAGTGAGGTCCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAGAAGACAGTCCCAGACAGCCTCCCCAG
CTTCTGCCCGTCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACG
ATCGAGAACTGCAGAGACTCACCTGATTGAGGATCAGACCCCTCCAGGCAAGGAGAAGTCA
GAGGCTGATTCTGTGAGAAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHL LGDPHTKNTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLES GCFQNLTCSPWACEQGTTPPMISWIGTSVSLDPSTTRS
SVLTLPQPQDHGTS LTCQVTFPGASVTTNKT VHLNVSYPQNLTMTVFGDGTVSTVLGNSSSL
SLPEGQSLRLVCAVDVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVS LQSKATSGVTQGVVGAGATALVFLSFCVIFVVVRS CRKKSARPAAGVGDGTGIE
DANAVRGSASQGPLTEPWAEDSPDQPPASARS SVGEGELQYASLSFQMVKP WDSRGQ EATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACCTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
 CCBAAGGAAGTGCAGCTGAGACTCAGACAAGATTACAATGCAACCAACTCAGCTTCCTGCTGTTTC
 TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAGGAATGGACCTGT
 TCTTCGCTCCATCTCTGCCCAGAAGCTGCAAGGAATCAAGACGAATGTCTAGTGCATTTGA
 TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
 GGGGTGGCGCTGGACCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
 GGCGATCGCTGGTCCAGTCAGCAGGCGACGAAGCAGACTACCCAGAGGGGGACGGCACTGGGC
 CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
 ACGACATCCAGGCCAAGGACCTGGGCACTGGCACGTGCCCAATAAGTCCCCATGCAGCACTGG
 AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCTCCAGACACTGGGACATAATCT
 GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAGTGTGGACTGACAACGGCC
 CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
 GGCCAGCGGGAATTCACTGCGGATTTGTTCACTTCAGGTTATTTAATAACGAGAGACGCCAA
 CGCCTTGTGTCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
 GATACTTTCCAGAGCCAGTCCCAGCAGTGTGGAGATTTTCTGGTTTGATTGGAGTGGATAT
 GGAACCTATGTTGGTTACAGCAGCAGCGTGAGATAACTGAGGCGAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
 CAACTTACCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
 AAAAAA

FIGURE 88

MNQLSFLFLIATTRGMSTDEANTYFKEWTCSSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDWSSQQGSKADYPEGDGNWANYNTFGSAEAAAT
SDDYKNGPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYGY
EGKCDWDNGPVIPIVVYDFGDAQKTASYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTTCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCTCCGAGGGCCATGGGCCGGTCTCAGGGCTTGTCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCACTCACTTATTCTGGTCCCGGACAGCAACATACAGGCCCTG
CCTGCCTCTCACGTTCACTCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACTCCCTGGGCTCTTTGCAGTGGAGCTGGCCGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACTCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTCGAGCGTTGGGAGTGCATACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTTGGGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGATTGTATTAGTGCTTTGTATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTTAGGGACAATTAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGAGAGCGAGGGGACGAGAGCGGAGGAGAAATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTCTGGTGACGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCTCCCGGGTCAGAGGACCTTGACGCTGATGACCAGAGGGCCAGCCCCGGCCCCGGTGCCT
CGGAAGCGGGGCCACATCTCACTAAGTCCCGCCCATGGCCAAATCCACTCTCCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACGCCCAACACAGCCCCC
CACCTCAGCCAAAGTGAAGAAAATCTTGGCTGGGGCAGATTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTCCCCCAGTAAGCTGTAG
AGTTCCACAGGAACAGCAGATCTTCATCGAAGCCAAGGCTCCAAAATCTTCAACTGCCGATG
GAGTGGGAGAAGGTAGAACGGGGCCCGCGGACCTCGCTTTTGACCCACGACCCAGCCAAGATCTG
CTCCCCGAGACCAGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGACTATCGGCTGGTCCGAAGGTGTGCCAGATTACAATAC
CATAGTGATACCCCTACTACCATCTGGGTGACCCCGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCCTC
AGGCAAGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGCCCAAGTCTCAAGTGGCAG
AGAAGGGTCCCAAGTGTGTGCCAACCTGAAGCTGTGGAGTGACTAGATCAGAGGACACTGG
AGGAGGAGTGGGCTCTCTGTGACGCTCACAGGCTTTGCCACGAGCCACAGAGAGATCTGGG
TCCCCGAGGCTGTGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCAATGGGAGGAAGCTAAGC
CCTTGGTCTTGCCATCTTGAGGAAGATAGCAACAGGAGGGGGAGATTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTTAGGACCAAGTCAAGAGGTGGGGTGGG
GCCAGAGAGCTCTCCAGCCCCTAGTGTGGGCCCTTGAGCCCCCTGTGCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGGTCTTGTATGCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCAACCCATGTGTATG
GCACACCATCTTAAAGCTAAGACAGGACGATTGTGGTCTCTCCACACTAAGGCCACAGCCCATC
CGCGTGTGTGTGTCCTCTTCCACCCCAACCCCTGTGGTCTCTCTGGGAGCATCCATGTCCCG
GAGAGGGTCCCTCAACAGTCAGCCTCACCTGTACAGCCGGGGTCTCCGGATCTGGATGGCG
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGCGGGCCGACAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTGTGGGTGGGGGAGGGAGGGAAGTCTTGTGAACCGCTGATGTGCTACTTT
TGTGTGAAGATCGTGTCTTGGAGCAGGAATAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLLYLVICQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGPPNRPNNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVYIAFYSTDYRLVQKVC PDYNVHSDTPYYP SG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGGCTCGCC
CTTTATGTCCTTCACCATCGCCATCGAGCCGTTCGCTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCICTACTGATTTCGTCCTTTGTTGGTTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCCTGTCTATATCCAAGAA
ATGTTCCGATTTCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTCCCTTTGTGAATACCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCTAT
GGAGATTCTCCTCAATTCCTCTTTATTCAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGTATGGCTGTGAGAAGAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAGCTTTCTCTTTACRACCAGCGCTCCAGAT
AACCTCAGGGAACCGCACTTCCCAACCGCAGACTACACTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGGAAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIDNKG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGPTVGIHGDSPOFFLYSAFMTLVIIILLHVFVGIVFFDGEKKKKGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFIAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
 GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGTCTCTAGGATCAAC
 TOGGTCATTACACAGCTCAAACCTGCTTTGGGACTCCCTCCACAAAACCTGGCTCCGGATCAGG
 GAACACTACCAAACCAACAGCAGTCAAATCAGTCTTTCCTCTTTAAGTCTGATACCATTAACA
 CAGATGCTCAGACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
 CCAGACCCACCCATTGACCCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTAC
 CAATTTTGTGCACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAATC
 TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCACCACTCAGGCAGGGGC
 TAATCCAGATGTCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
 GAACCCAGCAGGCGCCTCCCAACTCCAGTGGCACAGATGACGACTTTGCAGTGACCAACCCCT
 GCAGGCATCCAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAATGGAATTCA
 GTAAAGCTGTTTCAAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
 ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAGAGAAATTAATCTTAAAT
 TACCTGAAATATTTCTTGAATTTTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTAAAAA
 CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
 TAAACATATTTGGAAAACCTGGAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPAIGLPPTKLAPDQGLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLNPAAGMTPGTQTHPLTLGGLNVQQQLHHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLSLIHSLFPGGILPTSQAGANPDVQDGLPAGGAGVNPATQGTAGRLPTPSG
TDDDFAVTTTAGIQRSTHAIEEATESANGIQ

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
219

FIGURE 97

GCTCAAGTGCCTGCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCCTTGCTTGGCAGCTGGACCAAGGGAGCCAGCTCTTGGGCGCTGGAGGGCTGTCTTGACCATG
 GTCCCTGCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGCCACGCTGC
 AGAGCTGTCTGTGGAAGTCCAGAAACTATGGTGGAAATTTCCCTTTATACCTGACCAAGTTCG
 CGCTGCCCGTGAGGGGGCTGAAGGCCAGATCGTCTGTCAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATGCAGATTCTGGCTTCCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTCAACCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCAATTCTCTCAAGGCATCTAC
 AGAGCTCGGCTGAGCCGGGGTACCAGGCTGGCATCCCTTCCTCTTCTTGGAGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACTCGGATCTTGATCCACATCTTGAGCCAGGCTCCAGCCAGC
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAGGGGAGC
 ACCAGCCTTGACCAGCCCTGGAGAGGACCTACCAGCTGTTGGTGACAGGTCAAGGACATGGGTGA
 CCAGGCCCTCAGGCCACAGGCCACTGCCACCGTGAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCTTAGAGCCTATCCACCTGGCAGAGAATCTCAAGTCTCTATACCCGACCACTGGCCCAAGTA
 CACTGGAGTGGGGTGATGTGCATATCACCTGGAGAGCCATCCCCGGGACCCTTTGAAGTGAA
 TGCAGAGGGAACTCTTACGTGACCTAGAGAGCTGGACAGAGAGCCAGGCTGAGTACTCTGCTCC
 AGGTGCGGGCTCAGAAATCCCATGGCGAGGACTATGCGGCCCTCTTGAGAGTGCAGTGTCTGGTG
 ATGGATGAGAATGACAACGTGCTTATCTGGCCCTCCCGCTGACCCCAAGTCAAGCATCTCTGAGCT
 CAGTCCACCAAGTACTGAAGTGAAGTACTGTGAGAGGATGACAGTGCAGCCCGCTCCCCCA
 ATTTCCACGTTGTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTGAGAGGGAGAGCCTTC
 CAGGTGGACCCCACTTCAGCCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCAGACAT
 CCTGCTCTGTGCTGGCCATGGACTGGCAGGCCAGAGGGTGGCTTCAGCAGCAGCTGTGAAG
 TCSAAGTCGAGTCAAGATATCAATGATCACGCCCTGAGTCTATCACTTCCAGATTGGGGCT
 ATAGACTCCCTGAGGATGTGGAGCCCGGAGCTCTGGTGGCCATGCTAACACCCATTGATGCTGA
 CTTGAGCCCGGCTTCGCGCTCATGGATTTTGCGATTGAGAGGGGAGACACAGAGGCACTTTTG
 GCTCGATTGGGAGCCAGACTCTGGCATGTTAGACTCAAGCTCTCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCAAGGTTGGTGGTGGTGGTGCAGAGTTGGCGAAGCTGGTGGGGCCAGGCC
 AGGCCCTGGAGCCACCGCCACCGCTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCAAGTTGG
 ACCAGGAGGACTACGAGGCCAGTGTCCCATCAGTGGCCCAAGCCGGCTCTTCTCTGACCACT
 CAGCCCTCGGACCCCATCAGCCGAACCTCAGGTTCTCCTAGTCAATGACTCAGAGGGCTGGCT
 TCGCATTGAGAAATTTCTCCGGGAGGTGCACACCGCCAGTCCCTCGAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGTCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAAGACATGGCTTGATGCTGAGTGSACCCAGCAAGGACCCGATCT
 GGCCAGTGGGACCGGTCTCCATCAGCTTCAACCTTGGTCCCAACCCAGGCTGCACGGGATTGGC
 GCTCCAGACTCTCAATGGTTCCTACGCTTACCTTGGTCCCAACCCAGGCTGCACGGGATTGGC
 GAACACATAATCCCGTGGTGGTGCAGCAATGCCAGATGGGACGCTCTGCTGGTTCGAGTGTGAT
 CTGTGTGCTGTGCAAGCTGGAGGGGCAGTGCATGCGCAGGTTGGGCGCATGAAGGGCATGCCCA
 CGAAGCTGTGGCAGTGGGCATCTTTGAGGCACCTGGTAGCAATAGGAATCTTCTCATCTCTC
 ATTTTCAACCACTGGACATGTCAAGGAGAAGGACCCGATCAACACGACGACGCGTCCCT
 GAAGCGACTGTCTCAATGGCCAGGCAAGCTTACGTGGGAGCTTGGCTCTTGGCTCCATCTGAG
 TCCCTGGGAGAGGCCAGCACCCAAAGTCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCTCCGGGTGAGGCGACCATCACCATCACAGGCACTGTCTGCAGAGCCTGGACACCACT
 TATGGACTGCCATGGAGTGTCTCCAATGTCAAGGTGTTTGCCCAATATTAAGCCCCAGAGAA
 CTGGCTGGGCCCTATGGGAAAAA

FIGURE 98

MVPANLWLLCVSVFQALPKAQPAELSVVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRRLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPQPSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGQASGHQATATVEVSIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIYAAPLELHVL
VMDENDNVPICPPRDPVTSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPPEPDGVEGRA
FQVDFTSGSVTLGVLPFRAGONILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
FISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSEHVVVVQSVAKLVGPGGPGGATATVTLVERVMPKKLDQESYEASVPI SAPAGS FLTT
IQPSDPISRTLRFSLVNDSEGNLCIEKFSGEVHTAQSLQGAQPGDITYTLVEAQDTALT LAPVFS
QYLCTPRQDHGLIVSGPSKDPDLASGHGYPSTFLGPNPTVQRDWRLQTLNGSHAYLTLALHWVEP
REHII PVVSHNAQMWQLLVVRVIVCRCNVEGQCMRKVGRMKMPKLSAVGILVGT LVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

[illegible]

AGCTGACCGTGTCACTACAGTGGCTGGAGAAAGCTTAGGAAGAACCGAGCAATCCAGTGCGCCAGCGCTT
AGGTCACCAAGTTCTTCCAGAGCAAGCAAGCTAGGAGACCCAGCTCGTGGAGAGAGAGAGAGAGCTTTA
TCTCTTCACCTTCAGTCCCTTTCTCAAGATATCTTGTTCTTGCCCTCTAAGATCTTGGTAT
ATATAGGACCCAGGCAATCTTGCTTCCGSCACAAAGACAGATAGATGATGCAGAAAGAAATGAT
TCTCCTTATGTTTGGTCTACTTATGCTTTAGAGTGTGCACAAATTCCTAATGAGATCAGACC
TCTGCCAACACATGATGATGGCTGATCTCAGTGGAGCCAGCACAGCAACCACTCTGGGTGCCT
TGATCACTCCAGTGGGTGACGACAGCAACCACTCTCAGGTGCAGCGTGACCTTCAATGGGGT
GCATAGTGCACCAACTTCAGTCTCATACAACTCTCAGTGGATAGACAGCCCAACTCTTGAT
TTCAGCACAGCTGTGATGGGATCAGCATAGCCACCAACTCTCAGTGCACAGCAACCTCCAGTGT
GGCCAGACAGCCCAACTCTTGAGTCAGCACCAACTCTTGAGTGGGGCCAGCACAGTCAAGCTG
CTGGGTCTCAGTGTCCCTCAGTGGAGCCAGCACTGCCAACACTCTTGAGTCTCAGCACAGTCTCC
ATGAGGGCCAGCACTGCCCAACTCTTGATCTAGCACACTCTCAGTGGGGCCGACAGACAGCA
CAACTCTGACTTGCACCAACTCTCAGTGGGGTTCAGCACAGCCCAACTCTTGAGTGCACAGCA
CTCCTGAGTGGGGCCAGCACAGCCCAACTCTTGATGCACAGACTGTCCAGTAGGGCCAGCACT
GCCCAACTCTCTGAGTGCACCAACTCTCAGTGGGGCCAGCACAGCAACTCTTGAGTGCAG
AAGCACTCTCAATGGGGTGGCAGCAGCAACCACTCTGAGTCAGCAGCACTCTGAGTGGGGC
GCACAGCAACCACTCTGAGTCAGCACAGTGTCCAGTGGGGCCAGCACTGCCCAACTCTTGAT
TTCAGCACAGCTCTCAGTGGGGCCAGCACAGCAACTCTTGAGTGCACAGCAACCTCCAGTGG
GGCTAGCCAGCCCACTCTGAGTTCAGCACCAACTCTCAGTGGGGCCGACAGCAACCACTCT
TGAGTGCAGCAGACTGTCTGTTGATCAGCACAGTCACTTGTAGTCTCAGCACCACTCTCC
ATGAGGGGCCACACAGCCCACTCTTGATCTCAGCACAGCTCTCAGTGGGGCCACAGCCCACT
CAACTCTTGAGTGCACAGCACTGTCTGAGTGGGGCCAGCACATGCCAACCTCTTGAGTGCAGCA
CTCCAGTGGGGTTCAGCACAGCCCAACTCTTGAGTGCACAGCAACTCTTGAGTGGGGTTCAGCA
GCCCAACTCTCTGACTCAGCACAACTCTCAGTAGGCGCAGCACAGCCCACTCTGGGCTTAG
CAGTGTCTCAGTGGGATGCCAGTCACTCAACTCTTGAGTGCAGCACAACTCTCAGTGGGGCC
ACACAGCAACCACTCTGGTGTCCAGTGTCTCAGTCCAGCTCTGCAACAGCTCTTGATGTGT
ATGCACACACTCTCCATGATGCATCTCTGAGTGTGGGCGAAGCGCTGGTGGGTCTTGCTGG
GTGCTGGGGAATCTTCTCATCAGCTGGTCTGGTGTGGGCGCTGGGGCTCTTGTGGGCT
TCTCTCTTGTGTGGAAGAACAGCTCTGCTCGTGAAGAACCTTTAGACAGCTGTCTAACACCTCT
CATGGCTCTCAACATGGCTTGGTCTAGGCCCTGTGGAGGAATCTTAAAGCCGCCCAACAGGCCAG
GTGGAAGTCTCACTGTGTTCTGGAGGAGCCAGGATATCATGATAGCTATGGATAGAGCGGGAGGA
AGCAGGGCCGCTCAGAGCCCGGAGAGCAAGTCCGCAATCTTTCAAGGAAGAGAGACTTGGGCA
CCCAAGACTGCTTTTCTTTCACTTCAGCCAGAGAGCCCTCCAGCTTTGTTTGGATCTCTGAA
AATCTTGAAGAGGATATCTCACTTTCTTGCTTTTACAGCACTGGAAGAGAGATATATAT
TGCTCATTTAGGTAGAAGATATATCATCTTATACAGCAAGGAGAGAGCTGTGCTTGC
CCCGGGGTGGGATCTAGCTCTAGATGAATCTATGAGTATAGGAGAAAACTCCATGCTGCACT
CATCTGGCATCAAACTCCACAGTAAAACTCAAGACCTCAAAAAAATAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MMQMKGNVLLMFGLLLHLEAATNSNETSTSTANTGSSVSSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSVINSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNNGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSGASTATNSESSTVSSGISTVTNSES
STTSSGANTATNGSSSVTSAGSGTAALTMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLEAGLFFCVRNSLSLRNTFNTAVYHPRGLNHGLGPGPGNHGAPHRPRWSPNWFWRFPVSSI
AMEMSGGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCGGACGCCTCCGCGTTACGGGATGAATTACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCAGCTTGCCACGACCCCACTCGGCGTCGGCGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGGAACTATCAGGCTGAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGAAATTGTGTGCTTCTGAACCCATGTCATATT
AACGAGCGAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAAATCAT
GGTGTATCGAAAAGGATTACTTTTATCTACTGCTGTTTTGGGAAGCTTTTTTGAAGCATTT
TCATGCTGAGTCCCTTTTACCTTTGATGTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGAAAAAT
GATTATAACTGGGGATGCATTTGTTCCTGGAGAAAGAAGTGTCTATTATCATGAACCATCGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGTACCTCAGATTGGAGAAATTT
TGCTCACAAGCGAGTCTCAAAGGTGTTCTGTGATTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTGGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCACTCCTCATATTTCCAGAAGGACTGATCTCAGAGAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCAGTCCACCGGTATCCATAGACACCTCCCACATCCAAGGAGGA
CCTTCAACTCTGGTCCCAAAACGTTGGGAAGAGAAGAAGAGAGCTGCGTTCTTCTATCAAG
GGGAGAAGAATTTTATTTTACCGGACAGAGTGTATTCCACCTTGCAAGTCTGAACCTCAGGCTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCTGTCAGCCCTGCAATGTGCTACTCAT
ATATTTGTACAGTCTTGTAAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGTGGACTGGAGATCATAGAACTTGCAATGTACCAGCTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAGCTCAGATGCATTTTGCATGACTATGTCGAATATTTCTTACT
GCCATCATTTATTTGTTAAAGATATTTTGCACTTATTTTGTGGGAAAAATATTGCTACAAATTTT
TTTAACTCTGAATGTAATTTTCGATACTGTGTACATAGCAGGAGTGATCGGGGTGAATAACTT
GGGCAGAAATATTATAACAATCATCAGGCTTTTAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSPGIMVSWKGIYFILILFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVITGDADFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNFAEKNGLKQYEVVLPRTTGTFFVVDRLREGKNLDAVHDIIVA
YPHNIPQSEKHLQLGGDFPREIHFHVHRYPIDTLPTSKEIDLQWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVLVLLSLYWTLSFSPAMCLLIYLSLVKWFIIITIVIFVLQERIFPG
LEIIEELACYRLHLKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCAGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCATCCACCCTGCTGTATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGGTGTCAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCTTGGTGGGGGAGGACGCGTGTCTCTGCTCCCTCT
TTCTGAGACAGTGCAGAGGCTATGGAAGTGCGGTCTCTCAGGAATCAGTTCATGCTGTGGTC
CACCTCTACACAGATGGGGAACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAGGGGGCGGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCTGTATGGGTGCTGGTTCAGTTCACGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTTCGCTGTCTCAGGCTGGTTCCTCCAGCCACAGCCAAAGTGGAAAGGTCCACAAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGGAAATCCAGGTATTGATAGGAGAGACGTTTTTCAGCCCTCACCTTGGCGCTGGCTCTA
TTTTACTCGGGTACTCTGTGGTGCCCTGTGTGTTGTGTCATGGGGATGATAATTGTTTCTTC
AAATCCAAAGGGAAAAATCCAGGCGGAATCTGACTGGAGAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCACTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
TTTCTGATCTGAAAACGTGAACCATAGAAAAGTCCCCAGGAGGTGCCTCACTCTGAGAAAGAGA
TTTACAAAGGAAGAGTGTGGTCTCTCAGGGTTTCCAAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGACA
ATGTGACTTTGTCTCCCAACAATGGGTATGGGTCTCAGACTGACACAGAACATTTGTATTCTC
ACATTCAATCCCATTTTATCAGCCTCCCCCCAGCACCCCTCTACACGAGTAGGGGTTTCTT
GGACTATGAGGGTGGGACATCTCTTCTCAATACAAATGACAGTCCCTTATTTATACCTGTC
TGACATCTCAGTTTGAAGGCTTGTGAGACCTTATATCCAGCATCGATGTATGACGAGGAAGAG
GGGACTCCCATATTATATGTCCAGTGTCTGGGGATGACAGACAGAAAGACCTGTAAAGGGC
CCCCACACACAGCCAGACACAGCCAGGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
CGGAGCCTGCGCAGAGAGAGTCAAGCCCCCACTCTCTTTAGGGAGCTGAGGTCTCTCTGCC
TGAGCCTCGACAGCGCGAGTCACAGCTTCCAGATGAGGGGGGATTGSCCTGACCTGTGGGAG
TCAGAACCTTGGCTGCGCTGAAGTGGGGACGAATAGACTACATTAGTTTAGTTTGTGAAA
CTCCATCCAGCTAAGCATCTTGAACAAGTCACAACTCCAGGCTCCTCATTTGCTAGTACGGG
ACAGTGAATCTGCTCCACAGGTGAAGATTAAGAGACAAAGATGTGAATCATCTTCAGGTT
TGAGGACAGAGTGTTCCTAATGATGTGTTTTATATATACATTTTCCACCATAAACTCTGTT
TGTCTTTCCACATTAAATTTACTTTTCTATATCAAAATCACCATGGAAATGTTATGTGAACCC
TGCCTTGTGAGGCTCAAGAAATAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCACGAT
TACCTGTATCCAAAACAGGCAAGAAAACAGAAAGAGGAGGAAACATCAGGCTCATATTC
CTCATTAACACAGACACAAAAATCTAAATAAAATTTTAAACAATTAACATAACATATAATTA
AAGATGATATATAACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAAATATTTAAAT
ATCAACAGTGTAAATCAGACATTAAATAAGTAAATAAGAAAACCATTAATAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSABAMEVRFRNQPHAVVH
LYRDGEDWESKQMPQYRGRTFEFVKDSIAGGRVSLRLKNITPSDIGYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCCALCGVVMGMIIVFFK
SKGKIQAELDNRKKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTTHRKAPOEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPSTPPTTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLAPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTGGCAATGATGTATGGCCAGATGTGGTGAGGCTAGGAAAAGAG
 TTTGTGGGAACCTGGGTATCGGCTCGTCACTTCATATCCCTGATGTCTTGGCAGTGTGCATTGGA
 CTCACTGTTTCATTATGTGAGATATAATCAAAAGACCTACAATTTACTATAGCACATTTGCTATTTACAC
 TGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAAATTTTACAGAAATGAGCCAGAGACTTGANT
 CAATGGTGAAAATGCAATTTTATAAATCTCCATTAAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTTGTAGATTTTCACTCTACTGAGGATCCTGA
 AACTGTAGATAAAATGTTCAACTTGTTTTACATGAJAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
 ATCCTCACTCASTTAAAATTAATAAATCAACAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
 ACACGANGANGTAAAACCTTAGGTCAAGTCTCAGGATCGTTGGTGGGACGAAAGTAGAAGAGGGTGAATG
 GCGCTGGCAGGCTAGCCTGAGTGGGATGGAGTCACTCGCTGGGAGCAACCTTAATTAATGCCACATGSC
 TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCTGCCAGATGGACTGCTTCTTTGGAGTAACA
 ATAAACCTTCGANAATGAACGGGGTCTCCGAGAGAAATTTGTCCATGAJAAATACAACACCCATCACA
 TGACTATGATATTTCTCTTGCAGAGCTTTCTAGCCCTGTTCCCTACAGAAATGCAGTACATAGAGTTTGT
 TCCTGTGATCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAJAAAT
 GATGGTTACAGTCAAAATCATCTTCGACAGCACAGTGACTCTCATAGAGGCTACAACCTGCAATGAACC
 TCAAGTTTACAATGACGCCATACTCCTAGAATGTTATGTGCTGGCTCCTTAGAGGAAJAAACAGATGCA
 GCCAGGGTGACTCTGGAGGACCCTGGTTAATTCAGATGCTAGAGATATCTGGTACCTTGCCTGAATAGTG
 AGCTGGGAGATGAATGTGGGAACCCAAACAGCCTGGTGTATATACTAGAGTTACGGCCTTGGCGGAGTG
 GATTACTTCAAAAACCTGGTATCTAAGAGACAAAGCCCTCATGGAACAGATAACATTTTTTTGTGTTTTG
 GGTGTGGAGCCATTTTAGAGATACAGAAATGGAGAAGACTGCAAAACAGCTAGATTGACTGATCTCA
 ATAAACTGTTGCTTGATGATGATATTTCTTCCAGCTCTGTTCCGACAGTAAAGCATCTGCTCTGCOCA
 GATCAACTCTGTCACTGTGAGCAATAGTGAACCTTTATGTACATAGAGAAATAGATAATACATATTAC
 ATTACAGCCTGATTCATTGTTGCTCTAGAAGTTTGTCAAGAAATTTGACTTGTGTACATATAATTTGTAA
 GCATATATACAATTTGAAGCACTCCTTTCTCAGTTCCTCAGCTCCTCTCATTTACAGCAATATCCATTT
 TCAAGGTGCAGAACAGGAGTGAAGAAATATAAGAGAAAAATCCCTTACATTTTATTTGGCACAGAA
 AAGTATTAGTGTTTCTTAGTGGAAATATTAGAAATGATCATATTCAITTAAGAAAGTCAAGCAAGACACA
 GCAGATACCAATCACTTCATCATTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAGCCAG
 ATATATCCTTATTTTCATTTCACAAACACTACTATGATAAATGTGAAGAAAGATTCTGTTTTTTGTGACCT
 ATATAATTTATACAACTTCATGCAATGACTTGTCTAAGCAAATTAAGCAAAATATTATTTAAACATTG
 TTACTGAGGATGTCAACATATAACATAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYNSTLSFTTDKLY
AEFGREASNNPTEMSQRLESMVKNAFYKSPLEEFVKSQVKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMEFTGTGALKNDGYS
QNHRLQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGFPLVSSDARDIINYL
IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCGCGCTCTCCGCGAAGAAGTCCCTG
 CCCCAGTAGCCCCCGCGTGGCTCCCGACTATCCCAGCGGGCGTGGGACACGGGCCAGG
 GCCAGAGATCGTGGCGTTTGGCCCTGGGAGTAGGATGTGGTGAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACATGCGCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCGTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
 ATAATGTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCCT
 GTGGTTCATCGGTGATGATTGCTGTTTGGCTGTTTTCCTTATCATTGTGGGATGTAGGATATTG
 TGGAACGTTGAAGAATCTGTGCTTCTTGCATGGTACTTTGGAAGTTTGTCTGTCACTTTCT
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCAGTACAAATGGTCA
 GATATGTCACCTTTGAAAGCCAGGATGACAAATATAGGATTACCTAGATATCGTGGCTTACTCA
 TGCTTGGAAATTTTTTCAGAGAGATTAAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTAGAGAATCCCAGGATGTCCAAA
 CAGGCCACAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGGGAAGAAATGTATTCCCTT
 TTTGAGAGAACCAACAACTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAATGATGTCTTGAAGAATGACAACTCTCAGCAGCTGTCTGCTCAGTAGAACT
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAACAGCTTTAATACAACT
 TTGAGATGGAGAGTTTAAAGAAATGTGCACAGAAGAAACCAACAACTTGTTTATTGGACT
 TGTGAATTTTGTGATACATCTATGTGTTTCAGAAATATGTAGAAATAAAAATGTGGCATAAAA
 TAACACCTTAAGCATATCTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCAATAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGGAAGCAGATGTCATACCCACTGTGTAGGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTGGAGCAGCATGGTTGATTAGCAATTCGCCA
 TCAGTGCACACAGTGCATATGTGTGGGACTGGAGCCATGTAAGGTTGATTACTTCTACCA
 CTAGATATAAAGTACTAATTAATGCTAACATAGGAGATTAGAAAATCTAATACTTTTATTA
 CTCAGGCATCTATTCTCTGATGCTAAATAAATATATATCAGAAAACCTTTCAATATTGGTGACT
 ACCTAAATGTGATTTTGTGCTGCTACTAAATATTTCTACCACTTAAAGAGCAAGCTACACAT
 TGTCTTAAGCTGATCAGGATTTTTGTATATAAGCTCTGTGTTAACTCTGTAATTTCACTGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATATGAAAAGAAAATTTGCTGCTATAGCATCATTT
 ATTTTATGCTTTTCTGTTAATAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTATTAAATACTTAACCACTAATTTGAAATTAACAGTGTGATACATAGGAATCATATTTC
 AGAATGTAGTCTGGCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTCAGA
 AAGGACTGTATGCTGTTTTCTCCCAATGAAGACTTTTTTGACACTAAACACTTTTTAAAAA
 GCTTATCTTGGCTTCTCCAAACAGGAAGCAATAGTCTCCAAGTCAATATAAATTTCTACAGAAAA
 TAGTGTCTTTTTTCCAGAAAAATGCTTGAGAAATCAATAAACATGTGACAAATTTAGAGATT
 CTTGTTTATTTCAGTGATTAAATACTGTGGCAAAATACACAGATTATTAATTTTTTTTACAA
 GAGTATAGTATATTATTGAAATGGGAAAAGTGCAATTTACTGTATTTTGTGATTTTGTGTTAT
 TTCTCAGAATATGGAAGAAAATTAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAMRDYLNNVLTLTAEATRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWITYEQELMVPVQWSDMVT
LKARMTNYGLFPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTENDWPPDSCCVREFFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLAGTKQLQVLRPLGISIGVTQILAMILTITLLWALYYDRREP GTDQM
MSLKNDNSQHLSCPSVELLKPSLSRI FEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCACTCATCCTCATCCTCTTCCCTGTATAAAGCCCCTACCGATGCT
 GATAAAGCTTTCTCGTGAGAGCCTAGAGGCCCTAAAAAAGAGCTTTGAAGAAGAGGGGCAAAAGGAACA
 CCAGATTAAAGAGATTTCAGTGTCTTGGCAGTTGGTCCAGAGGATGCTCCATTCTGCTTCTCACTTG
 CCTCTTCATCACAGGCACCTCCGTGTCAACCGTGGCCCTAGATCCTTGTCTGTTACATCAGCCTGAATGAGC
 CCTGGAGGAACACTGACCAACAGTTGGATGAGTCTCAAGCTCCTCCTCTATGTGACAAACATGTGAATGGGGAG
 TGGTACCACTTCACGGCATGGGGGAGATGCCATGCCACTCTTCTGCATACCAGAAAACCACTGTGGAAACCCA
 CGCACTGTCTGGCTCAATGGCAGCCACCCCTAGAGAGCGACGGCATTGTGCAACGCCAGGCTTGTCCAGCT
 TCAATGGGAATGCTGTCTGTGGAACACACGGTGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
 CTGACCAAGCCAGGCTGTCTTCCAGCTTACTGTGCTCATTTTATGACATCTGGACGAGGACTGCCATGG
 CAGCTGCTCAGATACCAAGGATGCATGCGCTCCAGGAATGTGCTAGCCCTGACAGGCAAGATGCTTTG
 ATGAATAAGATTGACCAAAACAGGCTGCTGCACTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT
 GAGTGTGGCTTGGCCGTGTCTAAGAAATGATGGCAAGCTTGTGAAGAGCTTGAAGATGCCACAATAACAA
 TGGTGGCTGCAAGCACTTTGCTTGGATCTGAGAAGGCTACCACTGTGAATGTCCCCGGGGCTGTGCTGT
 CTGAGATAACCACTTGGCAAGTCCCTGTGTGTGCAAAATCAATGCCATTGAATGAACATCCCAAGGAG
 CTGCTGTGTGGCTGGAGCTTCTTCTGACCAACACCTCTCCGGGAGAGTGTCCAAGGCCACCATGTCAACAT
 CCTCTTCTCTCAAGACATGTGGTACAGTGGTGATGTGGTGAATGACAGATTGTGGCCAGCAACCTGTGSA
 CAGGTCTACCCAAGCAGACCCCGGGAGCAGCGGGGACTTCATCATCCGAACCAAGAGTGTGTATCCCGGTG
 ACCTCCGAGTTTCCACGCTGTGACACCATTTCTGAAGGATAGCTTCCCAACCTTCGAAACTCCCACTGGAAT
 CATGAGCGGAATCATGGATCTTCCATTCACTCTGGAGATCTTCAAGGCAATGAGTTTGAAGACCTTACC
 GGAAGCTCTGCCACCTCAAGCTTCTGACTCCCTTACTTTGGCAATTGAGCCCGTGTGTCAGCTGAGCGGC
 TTGGAAGCTTGGTGAGAGCTGCTTTGCCACCCCACTCCCAAGATGCAGAGGTCTGGAATACTAGCTCAT
 CCGGATGGCTGTGTTCAGATGACTCGGTAAAGCAATACATCCCGGGAATCACTAGCAAAAGCACTTCCAGG
 TCCCTGTCTTCAAGTTTGTGGCAAGACCAAGGAAGTGTTCCTGCACTGCGGGTCTTGTGTGTGGAGTG
 TTGGACGAGGCTTCCCGTGTGCCAGGGTTGCCACCGCGGAATGCTGCTGGGCGAGGAGGAGGACTCAGC
 CGGTCTACAGGCCAGACCTAACAGGGCGGCCGATCCGATCGACTGGGAGCACTGCTGCTAGGCCATACCTC
 GAGTCCCTGCATTGGAGCGCTGTCTTTGGAGCTTCTCCGCCACCGCCCTCAAGAACATCTGCCAACAGC
 TGGGTTCAAGCTTCACTGTGAGTTCAGACTCCAGCAACCACTCACTCTGATTTGCTGCATTCAATGAGGGA
 CAGGTCAACAGCACTGCTGAACAATGTGGCTGGTGGGTTTCACTTCTTAGGGTGAAGCACTAAACTGTCCA
 CCGAGAAGCACTCACCCATTTCCTCATTTCTTCTTCACTAAATACCTGCTGTATGCTGCAATCAGAC
 CACAAAATCAGAACTGGGTATATATTTCAAGTTACAAACCTAGAAAAATTAACAGTTACTGAATATATGA
 CTTAAATACCAATGACTCCTTAAATATGTAATATAGTTATACCTTGAATTTCAATTCAAATGACAGATAA
 TTATACGAATTTGGAAGTGTATCAATAAACACATATATAATTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHLQDLSQGPFPLCDNHVNGEWYHFTGMAGDAMP
TCIPENHCGETHAPVNLGSHPLEGDIQVRQACASFNQNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFVYDIEDCHGSCSDTSECTCAPGTVLGPRQTCFDENECEQNGGCSEICVNLKNSYRCECGVGRV
LRSDGRTCEDVEGCHNNNGGCSHCLGSEKGYQCECPRGLVLSEDNHTCQVPELCKSNAIEVNI PRELVGG
LELFLINTSCRGVSNQTHVNIILFSLKTCGTVVDVVNDKIVASNLVTGLPKQTPGSSGDPFI RTSKLLIPVT
CEFPRLTYISEGYVPNLNRSPLNLSRNHGIFFPFTLEIFKDNFEZEPYREALPTLKLADSLYFGIEPVVHV
SGLESLSVESCFATPTSKIDEVLYLIRDGCVSDSVKQYTSRDHLAKHFQVPVFKFVGKDKHEVELHCRV
LVCGVLDERSRCAQCGHRRMRGAGGEDSAGLQGGTLTGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZF domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAGGATGCTGGGCGTAGGGACCAAGGCTTGCCCTGCACCTCGG
 GCCTCCTCCAGCCAGTGCTGACCAAGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
 CCTGCTGCTTGGGGTGACAACTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCAGAGGCCAGCAGTGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTGATGATGTCRAACCCCTGCGCAAAACCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTTGT
 GGTGTCTCATCAAGGTGATTCTGGATAAACTACTTCTCTCTCGGGGAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCGAAGGGGCTGAGTGAGGCTCCGCTCTCCAGGAGCGATCCACACTGCAGTGCTGGAGCTCGGCCAC
 AGGGAACCTGGTCTCTGCTGTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCGAGTGG
 GCTACAGCAGAGCTGTGGAGATTGGCCAGACAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGCATGCGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCATGCTTGTGCTGT
 TGGGAAGAGCCTGAAGACCCCGCTGTGGTGGGGAGGAGGCCCTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTAGCAAAACAGCACGCTGTGTGGAGGGAGCATCTGGACCCCCACTGGGTCTCACGGCA
 GOCCTGCTTCAGGAACATACCGATGTGTTCACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCGAG
 CTTCOCATCCTGGCTGTGGCGAAGATCATCATATTGAATTCACACCCCATGTACCCCAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCTCACTTCTCAGGCAAGTCAAGGCCATCTGTCTGCCCTCTTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGAGGGAA
 GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCAGCGTGCAATGCGAGCGATGCGT
 ACCAGGGGGAGTACCGAGAGATGATGTGTGAGGCAATCCCGGAAGGGGGTGTGGAACACTGCCAGGGT
 GACAGTGTGGGCCCTGATGTACCAATCTGACCAAGTGCAGTGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGGGCCAGCACCCAGGAGTATACACCAAGTCTCAGCCTATCTCAACTGGATCTCAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAAGTGTGGGAGCGCTTCTCTCTGCCCTGCCACCT
 GGGGATCCCCAAGCTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTTGGAGCAGCAAGGGCCTCAATCTCTGTAAGAGACCTGCGACCCAGAGGGCGCCAGAGGAGATCA
 GCAGCCCTAGCTGGGCCACACTTGGTGTCTCCAGCATCCAGGGAGAGACAGCCCACTGAACAAAGGTCT
 CAGGGGTATGTGTAAGCCAAAGAGGAATTTCCCACTACTGATGGAAGCAGGCTGTCTTGTAAAGGCC
 CAGATCACTGTGGCTGGAGAGGAGAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCAACCATCCCCAA
 GCCTACTAGAGCAAGAAACAGTTGTAAATATAAATGCACTGCCCTACTGTGGTATGACTACCGTAACT
 ACTGTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGAACATCTCTGSCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRI PMETFRKVGIP III IALLSLASIIIVVVLKIVLDKYYFLCG
QPLHFIFRKQLCDGELDCPLGEDEEHCVKSFPEGFAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGGEEASVDSWFWQVSIQYDKQHVCGGSILDPHWLTAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQDSSGGPLMYQS
DQWHVVGIVSWGYGCGGFPSTPGVYTKVSAYLNWIYNVWKAEI

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACCTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCTACTCTCTACCTACATTAAAAATC
TGTTTTTTTGTCTCTTTGTAAGTACGCTTTACCTTCTTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTACCCACACCGTCCCCCTCGAAG
CCGGGGACAGCCTCACCTTGTGGCTCTCGCTGGAGCAGTGCCCTCACCACTGTCTCAGCTCT
GGAGGCAGTACTCGGGCAGTGCAGGTAGCTGAGCCTCTTTGGTAGCTGCGGCTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGATTGACAAGCCGAAGATTTCATAGGCGATGGCTCCCACTGCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACTGTCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCCATCTAACCTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAACCTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCAATCTCTTGGAACTAGAGG
GAACGCCGAGGAAAGCAAAGTGGCAGGGAAGAACTTGTGCCAAATTATGGGTGAGAAAGATG
GAGGTGTTGGGTATCACAAGGCATCGAGTCTCCTGCATTGAGTGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTTCGCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCGCTCTGGGAAGGAGCAAGTGACCAATTTCCTCTCCC
CTCCTTCCCTCTGAGAGCCCTCCTATGTCCCTACTAAGCCACAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTGAGCAAGGCCCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAATGCTCCAGTAAGCACAGGCTGCAAAATCCCAGGCAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGCCCTTATACCCAGGAGACTTTGATTGAAATTTGAAACCCCAAAATCCA
AACCTAAGAACAGGTGCATTAAAGAACTAGTTATTGCCCCGGTGTGGTGGCCTGTAAATGCCAACAT
TTTGGAGGGCCGAGGCGGGTAGATCACTGAGTCAAGGCTCAAGACAGCCTGGCCAACTAGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCACTGGTGGTGTGCTGTATC
CCAGCTACTCGGAGGCTGAGACAGGAGATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTACGCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAAATTA
TGTTATTGTGTA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLNLWPKPDHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2

FIGURE 115

CAGCAGTGGTCTCTCAGTCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
 AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
 ATATGTAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCTCT
 GTTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
 TCTACGCAATGGAGAGAAGAAGAGATTACATGGNAATTGATCCTGTGACCAGAAGTAAATA
 TTCAGAAGCGGAAATGSCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGATACACTGG
 CATCTACTTCGTGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTGATTCTCTGAATTTT
 CTGAACCCAGAAGAGAAATAGATGAGAATGAAGAAATTACCACAACCTTCTTTGAACAGTCAGTG
 ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATCCAAAATTCGGA
 GATTGTGTATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAG
 ACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
 GAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
 AGAATCTCAATAAATGACTATACTGAAAATGGAATAGAATTGATCCATGCTGGATGAGAGAG
 GTTATTGTTGATTTTACTGCCGTGAGGCAACCGCTATTGCCGCCGCTGTGTGAACCTTTACTA
 GGCTACTACCCATATCCATCTGCTACCAAGGAGGACGAGTCATGTGCTGTGTCATCATGCCCTTG
 TAACGTGGTGGTGCCCGCATGCTGGGGAGGGTCTAAAGTAGGAGTTTGAGCTCAAATGCTTAAAC
 TGCTGSCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
 GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTCATGTTCTAATAAATCTTACA
 TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME
HTFYNSNGEKKKIYMEIDPVRTETIFRSGNGTDETLVHDFKNGYTGIFYVGLQKCFIKTQIKVIP
EFSEFEEIEIDNEEIIITTFEQSVIWWPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRNGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGAGCGCTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGCGCGCA
GGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTTCCTCCTGTCCATCTCTGGGGCT
GGCGCGCTGCATCGCGGCCACCGGATGACATGTGGAGCACCCAGGACCTGTACGACAAACCCCTGCACCT
CCGTGTTCCAGTACGAGGGGCTCTGAGGAGAGCTGCGTAGGGCAGAGTTGAGGCTTCACCGAATCGAGGCC
TATTTACACATCCTGGGACTTCCAGCCATGCTCGAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCT
GGTGCCATTGGCTCCTGGTATCCATCTTTGCCCTGAAATGCATCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACTGACACTGACCTCCGGGATCATGTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACTTCTGGATGTCCACAGCTAACTGTACACCGGCATGGGTGG
GATGTTGCAGACTGTTACAGACAGGTACACATTTGGTSCGGCTCTGTTCTGGGCTGGGTCTGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGGCTGGCACCAGAAAGAACCACTACATAAA
GCCGTTCCTTATCATGCTCAGGCCACAGTGTGCCTACAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGTCCAAACACCAAAACCAAGAGATATACGATGGAGGTGCCCGACAGAGGACGAGGTACAACTTTATC
CTTCCAAACAGACTATGTG**TAA**TGCTCTAAGACCTCTCAGCACGGCGGAAGAACTCCCGGAGAGCTCA
CCAAAAACCAAGGAGATCCCATCTAGATTCTCTTGTCTTTGACTCACAGCTGGAGTATGAAAGAGCT
CGATTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTCAATCCTCTATTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTAACTCTCTCTCACATTTGATGATTAGACAGACTCCGCCCTC
TTCTCCTAGTCAATAAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAACTTTTGAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGCAATAGAGAGAAGATATTGTANTCTCTCCAGCCCATGATCTCGGTTTCTCT
ACACTGTGATCTTAAAGTTACCAAAACAAAGTCATTTCAAGTTGAGGCAACCAAACTTTCTACTGCTGT
TTGACATCTCTTATTACAGCAACCACTTCTAGGAGTTCTCGAGCTCTCCACTGGAGTCTCTCTCTGT
CGCGGGTCAGAAATGTCTCCTAGATGAATGAGAAAATATTTTTTTTAAATTAAGTCTTAATATAGTTAA
AATAAATANTGTTTAGTAAATGATACACTATCTCTGGAATAGCCTCACCCCTACATGTGGATGAAGA
GAATGAAGAAATTAATGCTTTGACATTGTCTATATGGTACTTTGTAAGTCTATGCTTAAGTACAAATTC
ATGAAGCTCACACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAAGGATCACTTGAGCCAGAGGT
TCGAGACTAGCTGGGCAACATGGAGAAGCCTGTCTACAAAATACAGAGAGAAATACAGCCAGTCA
TGGTGGCATAACCTGTAGTCCAGCATTCGCGGAGGCTGAGGTGGAGGATCACTTGAGCCAGGGAGGT
TGGGCTGCAGTGAGCCATGATCACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
ARTAAAAATAAATAATGAACACAGCAAGTCTAGGAGTAGGTTAAACTAATCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLMRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSI FALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCTGCTTCAAAGCAGAAGTAGCASTTCCGGAGTCC
 AGCTGGCTAAAACATCATCCACAGAGGATAATGGCAACCAATGCGCTTAGAAATCGCTGGGCTGTTTCTTG
 GTGGTGTGGAAATGGTGGGCACAGTGGCTGTCACTGTCAATGCCTCAGTGGAGAGTGTGCGGCTTCATT
 GAAACACACATCGTGGTTTTTGAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
 CATCAGGATGCAGTGCAAAATCTATGATTTCCCTGCTGGCTTTTCTCCGGACCTACAGGCAGCCAGAG
 GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGCTTTTCATGATGGCCATCCTTGGCATGAAATGC
 ACCAGGTGCACGGGGGACAAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
 CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCTATA
 ACTCAATAGTGAATGTGCCCAAAACGTGAGCTTGGAGAACTCTCTACTTAGGATGGACCAGGCCA
 CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGTGCGTTTTTGTGCAACGAAAAGAGCAGTAGCTA
 CAGATACTCGATACTTCCCATCGCAACCCAAAAAGTTATCACACGGGAAAGAGTCACCGAGCG
 TCTACTCCAGAAAGTCAGTATGTGAGTTGTGTATGTTTTTAACTTTACTATAAAGCCATGCAAAATG
 ACAAATACTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTACTGTCTTAACTGCGCT
 AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCATATAAGCTATTTACAGCAGAAATGAGATA
 TTAACCCCAATGCTTGTATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
 CTCCTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
 ACGACATAGCATTAATGTACATAGATGAGTGTAACTTTATATCTCACATAGAGACATGCTTATATGGT
 TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAATAGAACTCAACTATTGCTTTTCAGGGAA
 ATCATGGATAGGGTTGAAGAAGGTACTATTATTTGTTAAAAACAGCTTAGGGATTAATGTCTCCCA
 TTTATAATGAAGATTAAATGAAGCTTTAATCAGCATTGTAAGGAAATGAATGGCTTTCTGATAT
 GCTGTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCTCTCTTCCAGAGGCTTTTTTTTT
 CTTGTGTATTAAATTAACATTTTAAAAACGAGATATTTGTCAAGGGGCTTGCATTCAAACGTCTT
 TTCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGTTTTAGGAAAGTG
 AAAATATTTTTGTTTTGTATTGAAGAAGAAATGATGCACTTTGACAAAGAAATCATATATGTATGGAT
 ATATTTTAAATAGTATTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
 TGTCTTGGTTTTTCATTGCTTACAAAAAACAAACAAAAAAGTTGTCCTTTGAGAAGCTTACCT
 GCTCCTATGTGGGTACCTGAGTCAAAATGTCAATTTTGTCTGTGAAAAATAAATTCCTCTCTTGTA
 CCATTTCTGTATTAGTTTTACTAAAATCTGTAAATACTGTATTTTCTGTTTATTCCAAATTTGATGAA
 ACTGACAAATCCAATTTGAAAGTTTGTGTGACGCTCTGTCACTTAAATGAATGTGTTCTATTGTCTT
 TATACATTATATTAATAAATGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTMPQMRVSAFIENNIIVVFENFW EGLWMNCVRQANIRMQCK
IYDSLALLSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRD FYNIVNVAQKRELGEALYLGWTTALVLI VGGALFCCVFCCKNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 121

GGAGAGAGGCGCGCGGTGAAAGGCGCATTGATGCAGCCTCGCGCGGCTCGGAGCGCGCGGAG
 CCAGACGCTGACCACGTTCTCTCTCTCGGTCTCTCCGCTCCAGCTCCGCGCTGCCCGGCAGCC
 GGGAGCCATGCGACCCAGGGCCCCGCGGCTCCCCGAGCGGCTCCGCGGCTCCTGTGCTCC
 TGCTGCTGCAGCTGCCGCGCCGTCGAGCGCCTTGAGATCCCCAAGGGGAAGCAAAAGCGCAG
 CTCGCGCAGAGGGAGTGGTGACCTGTATATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
 TGCTCGAGACGGAGCCCTGGGGCCAATGTTATTCGGGTACACCTGGGATCCAGGTCCGGATG
 GATTCAAAGGAGAAAAGGGGAATGTCTGAGGAAAGCTTTGAGGAGTCTGGACACCCAACATAC
 AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTGGGAAAATTCGGAGTGATCATT
 TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTGTTCAGTGGCTCACTTCGGCTAAAATGCA
 GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTCAGGACCTCTTCCC
 ATTGAAGCTATAATTTATTTGGACCAAGGAGCCCTGAAATGAATTCACAATTAATATTCATCG
 CACTCTCTCTGTGAAGGACTTTGTGAAGGAATGGTGCTGGATAGTGGATGTGTCTATCTGGG
 TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTCTCGCATC
 ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
 TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAGCAAAG
 CTAATATGTTTACAGACCAAAGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
 CTTCATCAAAAGTGGTTTCAATATTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
 CTCTCACCTATAATTTGGAATATTGTTGGTCTTTTGTCTCTTAGTATAGCATTTTTTA
 AAAAAATATAAAGCTACCAATCTTTGTACAATTTGTAATGTTAAGATTTTTTTTATATCTGT
 TAAATAAAATTTATTTCCACA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 122

MRPQGPAASPQRLRGLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQSPAGVPGR
DGSPGANVIPGTPGIPGRDGFKEGECLESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMNSTINIHRIS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCAGGGGAACCTGCTGGGCTATGGAAATCAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGGCTGGAGAATAACATCATGTTTTTCGATAAGGAAGAAATGTAGGATCCAGTTT'TTTTTTA
 ACCGCCCCCTCCCCACCCOCCAAAAAAGCTGTAAGATGCCAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGTGTTTTGCTGCGAATCGGGTGTGGGATTTATTTGTCTCTGGAG
 TGTCTCGTGGCTGGCAAGAAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAAATTTT
 TCTTCTCGGGTGCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCAATGCAACTG
 GCCCTTAAGCCAAAGCAAAGACCTAAGGAGCAGCTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCGGAACGAGGATGCCCTAAGGGCTGAGGTGTGAAGGCCAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCCTAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAACTTAAGTATAATCAATTTAAGGGCTCAACGAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATGACGAAATGCTTTTAATGGAATACGAGACTCAAGA
 GCTGATTCTTAGTTCCAAATAGATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAATTT
 TACGGAACCTGGATCTGTCTATAATCAGCTGCATCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGGTCTAAGCTCCCTGAGAACCATCCCTGTGCGAATATTCGA
 AGACTGCCGCAACCTGGAACTTTTGACCTGGGATATACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAGRACTTCACCTGGAGCAAAATCAATTTTCCAGCTCAAC
 CTGGCCCTTTTCCAAAGTTTGGTCAGCCTTCAGAACCTTTACTTTCAGTGGAAATAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTGTATTTATCAGGCAATGAGA
 TCGAAGCTTTTCAGTGGACCCAGTGT'TTCCAGTGTGTGCCGAATCTGCAGCGCTTCAGCACTGGA
 TCCCAACAGCTCACATTTATTGGTCAGAGATTTGGATTCTTGGATTCCTCAATGACATCAG
 TCTTGTCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTGTAAACTGGCTGAAAGTT
 TTAAGGTCTAAGGAGAAATCAATATCTGTGCCAGTCCCAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCACTGAAGAACTACAGCATCTGTGGCAAGATCTACAGAGAGCTTGTATCTGGCCAG
 GGCTCTCCCAAGCCGAGCTTTAAGCCCAAGCTCCCGAGCCGAGCATGAGAGCAACCCCTTT
 TGCCTCCGACGGTGGAGCCACAGAGCCCGGCCAGAGACCATGCTGAGCCGAGCACTCTCT
 TTCCAATAAATCATCGCGGCGAGCGTGGCGCTTTCTCTGTCCGTGCTGTCATCTCTGTGGTTAT
 CTACGTGTCTGGAAGCGGTACCTCGGAGCATGAAGCAGCTGCAGAGCGCTCCCTCATGGGA
 GGCACAGGAAAAAAGAAAGACAGTCCCTAAGCAAAATGACTCCGACCCAGGAATTTATGTA
 GATTATAACCCACCAACAGGAGACCGAGATGCTGCTGAATGGGAGGGGACCTGCACACT
 TAACAATCGGGCTCCAGGGAGTGTGAGSTATGACCATTTGTGATAAAAGAGCTCTTAAAGCT
 GGGAAATAGTGGTGTCTTATTGAATCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
 TCCCTCTCCCTCTCACTTTGGTGGCAAGATCTTCTTGTCCGTTTTTAGTGCATTCAATAACT
 GGTCAATTTCTCTCATACATAATCACCCATTTGAATTTAAATACCACATCAATGTGAAGCTT
 GAACCTCGGTTTAAATAATACTATTGTATAAGACCTTTACTGATTCCATTAATGTGCGCATTT
 GTTTTAAGATAAAATCTTTCATAGGTAAAAAATAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTLLTMLSSAERGCPKGCRCGKRVYCESQKLQEIIPSSISAGCLG
LSLRYSNLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRLNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNLSRLTIPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNALFPRLVSLQNLYLQWNKISVIGQTSWNTWSSLQRLDL
SGNEIEAFSGPSVFQCVENLQRLNLDNKNLTFIGQEILDSWISLNDISLAGNIWEC SRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLALALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRHRKKRQSLKQMTPTSTQEFYVDYKPTNTETSEM LNLGTGPCTYKNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCGGTCCCGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAATTCCTACTGTCCGGCTGCGCGGCTACCGTGCCCGAGCT
AGCAACCTTTCCCTGGATCTCACAAAACCTCGACTCCAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAGAGAATCTGCCCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACAGTAGT
GTATTCTGGAGGTGGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTATGGCCAGTTT
TTAGCCAATCCAACCTGACCTAGTGAAGGTTCAGATGCAATGGAAGGAAAAGGAACTGGAAGG
AAAACCATTCGATTTCTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCCAGCGATGTCA
TCAAAAGCAGAAATAATGAATCAACCACGAGATAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
ATCTTGGCTGAGAAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTORWPRASKFLLSGCAATVAELATFPDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLEVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSDCLIQAVQGEFGMSLYKGFLPSWLRMTPTWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

1-20
21-40
41-60
61-80
81-100
101-120
121-140
141-160
161-180
181-200
201-220
221-240
241-260
261-280
281-300
301-320
321-340
341-360
361-380
381-400
401-420
421-440
441-460
461-480
481-500

FIGURE 127

CGCGGATCGGACCCAAAGCAGGTTCGGCGGCGCGGCAGGAGCGGCGGCGGTGAGCTCCTCGAC
 CCCCGTGTGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCGCATGGGCCAGGCCGGCATGG
 AGCGGTGGCGGACCGGCTGGCGTGGTGACGGGGGCGTTCGGGGGGCATCGGCGCGCGCGTGGCC
 CGGGCCCTGGTCCAGCAGGACTGAAGGTGGTGGGCTGCGCCGCACTGTGGGCAACATCGAGGA
 GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGGACTTTGATCCCTACAGATGTGACCTAT
 CAAATGAAGAGACATCCTCCTCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
 TGCAATCAACATGCTGGCTTGGCCGGGCTGACACCTGCTCTCAGGCAGCACCAGTGGTTGGAA
 GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
 AGGAGCGGAATGTGGACGATGGGCACATTAACATCAATAGCATGCTTGGCCACCGAGTGTTA
 CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTAGCCGTCACTGCGCTGACAGAGGGACT
 GAGGCAAGAGCTTCGGGAGGCCACAGACCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
 AGACACAATTGCGCTTCAAATCCACGACAAGGACCTGAGAAGGCAGTGCCACCTATGAGCAA
 ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCGCACA
 CATCCAGATTGGAGACATCCAGATGAGGCCACGGGAGGCTGACCTAGTGACTGTGGGAGCTCC
 TCCTTCCCTCCCACTTCATGGCTTGCTCCTGCTCTGGATTTTAGGTGTTGATTTCTGGAT
 CACGGGATACCACTTCTTGTCCACACCCGACCGAGGGGCTAGAAAAATTTGTTGAGATTTTATA
 TCATCTTGTCAAATGCTTCAGTTGTAATGTGAAATGGGCTGGGGAAAGAGGTGGTGTCCC
 TAATTGTTTTACTTGTAACTTGTTCTTGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
 TCTTCCCTTTGACATGGGAAAGAGTGTGGCCAAATCCCATCTTCTTGCACTCAACGTCTG
 TGGCTCAGGGCTGGGCTGCGAGAGGAGGCCCTCACCTTATATCTGTGTTATCCAGGGCTCC
 AGACTTCTCCTCTGCTGCCCCACTGCACCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
 AGCCCACTTGGCTTCTGTGCTCCTCTGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAG
 CAGAACACCGAGGCTGGCCAGTGGATTTATGGTGATCATTAATAAAGAAAAATCGCAACCAA
 AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAARALVQQGLKVVGCARTVGNIEELARECKSAGYPGTLI
PYRCDLISNEEDILSMFSAIRSOHSGVDICINNAGLAREDTLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININSMGHRVLEPLSVTHFYSTATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTC TAC ATG GGC CTCTGCTGCTGGTGCTCTTCCTCAGCCTCTGCGGGTGGCCTACACCAT
CATGTCCCTCCCACCCCTCCTTTGACTGCGGGCCGTT CAGSTG CAGAGTCTCAGTTGCCCCGGGAGC
ACCTCCCCCTCCGAGGCAGTCTGCTCAGAGGGCCCTCGGCCAGAAATCCAGTTCTGTTTCATGC
CAGCCTGTAAAGGCCATGGAAC TTTGGSTGAATCACCAGATGCCAATTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGTCTGTTCTGTGCTGCGCTGTTTCATTT CAGTAGCCACAGCCACCTGTGG
CCGTTGAGTGCTTGAAA TGAG GGA CTGAGAAAATTAATTTCTCATGTATTTTCTCATTATTTA
TTAATTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAA TCAGGGTA ACTGGGATATCCATCACATCAACATTTATTTTTATCTTT
TTAGACAGAGTCTCACTCTGTCACCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCCTGCCAGGTTCAAGCGATTCTCATGCCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCACTAATTTTTGTATTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACCTCTGGCCTCAACAATCCACTTGCCCTCGGCCCTCCCAAGTGTTATGATTACA
GGCGTAGGCCACCGTGCTGGCCTAAACATTTATCTTTCTTTGTGTGGGAAC TTTGAAATAT
ACAATGAATTATTTGTTAACTGTCATCTCCCTGTCTGTGCTATGGAACTGGGACTTCTCCCTCT
ATCTAACTGTATATTTGTACAGTTAACCAACCTACTTCATCCCCACTCTCTATCTCTCC
AACCTCTGATCACCTCATTCTACTCTACCTCCATGAGATCCACTTTTTTAGCTCCGACATGTG
AGTAAGAAAATGCAATATTTGCTTTCTGTGCTGGCTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAATGACAGGATTCGTTCTTAATTTCAATTAAATAACCAACATG
GCAAAAA

FIGURE 130

MGLLLLVLFSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGGHTLGESMPFKRVCQDGNVRSFCVCAVHFSSHOPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAAACGGAAGCTACCTTGATATAAGACCTCAACACTGCTGACATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATTGGCGTGTTCCTTCAAGTAGACCTCTATCAGTTATGGCTAAATCCTG
 TOCATCTGTGTGCTGCGATCGCGGTTTCACTTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
 GAATACCGAGGATGCTACAACCTCTACTCTCAGAACCAACCAATAAATAATGCTGGGATTCCTTCAGAT
 TTGAAAACTTGTGAAAGTAGAAAATATACCTATACCCACACAGTTAGATGAATTTCTACCAACCT
 CCCAAGATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACCTTTCAA
 AAATTCCTATCTGGAGAATTACATTTAGATGACAACTCTGCTCTGCAAGTTAGCATAGAAAGGAGGACA
 TTCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTTT
 GCGCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTCTATCACCATCTCTCAAG
 GTCTCACTAGTCTAAAACGCTGGTTCTAGATGGAACCTGTTGAACATCATGGTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTTGACAGAGCTGCTCCCTGGTCCGGAATTCCTGACTGCTGCAOCAGTAAACCT
 TCCAGGACAAACCTGAGGAAGCTTTATCTTCAAGATAACACATCAATCGGGTCCCCCAATGCTTTTTT
 CTTATCTAAGGCAGCTCTATCGACTGGATATGTCGAATAAATCACTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACTGATTTCTCGCAACATCCCTGGTATTCGCGGTGCAAGATGAATG
 GGTACGTGACTGTTACAACTACTACTGTGAAGTCAACGTGCTGGGCTCATGTGCCAACGCCAGAAA
 AGGTTCTGGGATGGCTATTAGGATCTCAATGCAGAACTGTTGATTTGAGGACAGTGGGATGTATGC
 ACCATTCAATACCACTGCAATACCCAACACAGTGTATCTGCCCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAACAGCCAGATATTAAAGAACCCAGCTCACTAAGGATCAACAAACCACAGGAGTCCCTCAAGAAAA
 CAATTAACAATTACTGTGAAGTCTGTCACCTCTGATACCACTTATATCTCTTGAAGATTCGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGCATTTGGATCTATACAGAAACAAATGT
 AACAGGGGAACGAGTGAAGTACTTGGTCAACGCCCTGGAGCCTGATCAACCTATAAGTATGCATGGTTC
 CCATGGAACACGCAACCTCTACCTATTGATGAACCTGCTGTTGATTTGAGACTGAACCTGCACCCCTT
 CGAATGTACAAACCTACAAACCCCTCAATCGAGAGCAGAGAAAGAACCTTACAAAAACCCCAATTTACC
 TTTGGCTGCCATCATTGGTGGGGCTGTGGCCTGGTTACATTTGCCCTTCTGCTTTAGTGTGTGGTATG
 TTCTATAGGAATGGATCGCTCTTCTCAAGGAATCTGCAATATAGCAAGGGAGGAGAAAGGATGACTAT
 GCAGAGCTGGCACTAAGAGGACAACTCTATCTGGAATCAGGGAACCTCTTTTCAGATGTACCAAT
 AAGCAATGAACCACTCTCAAGGAGGAGTTGTAATACACCACTATTTCTCTTAATGGAATGAATCTGT
 ACAAAAACATACAGTGAAGCAGTAGTAACCGAAGCTACAGACAGTGTATTCAGACTCAGATCAC
 TCACACTCATGATGCTGAAGGACTCAGACAGACTGTGTGTTTGGGTTTTTAAACCTAAGGAGGTGATG
 GT

FIGURE 132

MISAASWIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFTYCNDRLFSTIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTTITYDSLSKIPLYL
EELHLDNSVSASVIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSLVRNGLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLNSLPQGIFFDDLNTIQLILRNPNWYCGCKMKWVRDNLQSLPV
KVNVRGLMCQAPKVRGMAIKDLNLAELFDCCKDSGIVSTIQITTAIPNTVYPAQGGWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALFMTALRLSLWLKLGHSAPFGSITET
IVTGERSEYLVTALEPDSPIYKVMVPMETSNLYLFDETPVCIETETAPLRMYNPTTTLNREQEKE
PYKNFNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKWNHSESSSNRSYRDSGIPDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTTCCAGAGTCTCTTGGCCAGGCCACCCAGGCTTCTTGGCA
 GCCCTGCCGGGCACATTGCTTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
 TGCAGAGGAGTCTGGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAAGCAGAG
 GCCGGCCATGGCCAGCGTGGGGTGTCTGCTCTTACTGACAGACTGCCACCGCTGTGGT
 CCTCTCACTGCCTGGGCTGGACACTGTGAAAGTAAAGCCACCAITGAGACCTGATCCTGTCT
 GCGCTGGAGAGGCCACCGTCTTCTAGAACAGAGGTGCTGAAATCAACCTGGATGGCATGGT
 GGGGGTCCGAGTGTGGAAAGCAGCTAAAGTGTCCGGGAGAAAGTGGGCCAGGAGCCCTGC
 TGCAGCCGCTGAGCTGCGCGTGGGGATGTGGGGGAGAACTGGAGGCTGCCATCCAGAGATCC
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCCTCAGCCCGG
 GTTTTGAAGCTCCACATGCTGGATCCACACTGATGCTCTTGGTGATCCCGCTTGGGC
 CCCAGGACTCATCTCAGAGGAGAGAGTGACGTGTGCTTGTGTCAGCTGTGGGAACCGGACG
 GACAGCAGCGAGCCCTGCGGCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
 AGGCTACTGCTGTCCACCAACTGCTCTTCTCTGCGCCAGAATGAGGGGATGCACACAGG
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCACATGATGGAATGAACGCG
 AGAGCTGAGGCCATCGGATACGCTACCCCTACCCGGGACATCTTCATGGAAACATCATGTTCTG
 TGAATGGCGGCTTCTCCAGTCTTACAAGCTCCGGTGGCTGGAGGCCATCTCAGCTGGCAGA
 AACAGCAGGAAGGATGCTTGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
 TATCAGCAGCATTTTTCGAGGAGAGTGAAGAGGCGAGAAAACAAATTCAGATTCTCGCTCTGT
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCTCTCGGTTCAAGC
 AATTCTCTTGCTCATCTCCGAGTAGCTGGGACTACAGAGAGCTGCCACCATACCTGGCTAAT
 TTTTATATTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGCTCGAATCTCTGAT
 CTCAGAGATCCGCCACCTCAGGCTCCCAAGTGTGGGATTATAGTGTGAGCCACCGTGTCTG
 GCTGAAAGCACCTTCAAAGAGACTGTGTTGAATAAAGGCCAAGGTTCTTGCCACCCAGCAGTCT
 ATGGGGGCTCTCTCCCTAGATGGCTGTCTCCCAACAACAGCCACAGCAGTGGCAGCCCTGG
 GTGGCTTCTATATCTTGGCAGAAATACCCCCAGCAACAGAGAGCCACCCATCCACACCG
 CACCAACCAAGCAGCGCTGAGACGGAGCGTTCCATGCCAGCTGCTGGAGAGGAGACAGACCC
 TTTAGTCTCATCCCTTAGATCTGGAGGGCACGGATCATCTCTGGGAAGAAGGCATCTGGAGG
 ATAAGCAAGGCCACCCGACACCAATCTTGAAGCCCTGATAGCAGGGCCAGGGTAGTGTGG
 GGCCGGGAGGACCCAGGTGTGAACGGATGAATAAAGTTCACCTGCACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERTVSSEQRPMASLGLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPILLQPLSLRV
GMLGEKLEAAIQRSLSLHYLKLSDPKYLREFQTLQPGFWKLPKAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSSPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLNARMRGCTQGPLQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQEGCGF
EPDAEDEELSKAIQYQGHFSRRVKREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCA**TAG**GGGGCCGCTCTGTGGGGCTCTTTCCCGTCTGCTGCTG
CTGCTGCTATCGGGGATGTCAGAGCTCGGAGGTGCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCCGCATAGGAGATCGCTTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGTGAAGC
CTCAGGACTGGATCTCGCGGGCCGAGTCTGGTAGACGGAGAAGAGCACGTCCGTTTCCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC
AGCTTACAGATTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTGAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTTAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCAATGGT
TATGATGATGGTCTTCCCTTTATTGATATTTGTGCTTCTGCCATAAGTGGTCAACACANGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAATCTAGCAGCGGCAGCAG
TAAACAGGCCAAAAGTGGGGCTGGCAAAGGAGG**TAG**TCAGGCCGTCCAGAGCTGGCATTTGCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGA AACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTAGCACATGTTTGTACTTGGTACACGAGAAAAACCGACTTTCATCTTTTGCTGTGAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCCATAGAAAAATGCCATTAATAAATTAT
ATGAACACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSGDVQSSEVPGAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGLKTDGSFVVHDIPSGSYVVEVWSPAYRFDVPRVDITSKGMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPVSEFMTRLFSSKSSGSSSGSKTKSGKGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCCAGCCACAGCTTCTGTGAGATTGATTCTCCCCAGTTCCCTGTGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCAAAACAA
GTTTGTACATTTCCCTGAAATGTCATTTCTATCTATTCACTGCAAGTGCCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTTACGGGAGGCTTGGCAGT
TTTTCTTACTCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCCTCAGC
CTTCTCTCTGTCGGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCTCGGAATCGATGCTGCCTCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
AACTACCAGACCCCTGACCATTATACTCTCGGAAGATCAGCAGCCTCGCCAATTCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGAACTAGACATTTCTTGCATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCCAGATGTATCTTATTTATGCATTACTTG
CTTCTTGCATGATTGTCTTTATGCATCCCAATCTTAATTGAGACCATACTGTGATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATAGTTAATATATTTATTTATTTTGTCTATTTA
ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGATTTTATACAGTAATAAAAAAACCTTGTAATTTCTAGAAGAGTGG
CTAGGGGGGTATTTCATTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGTAATTTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTGTGATGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCAATATTTGTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAGTTTTCTTGCAATACAAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGITCWALTAEPGWGQNGATTCA¹TNSHSDSEL²RPEIF
SSREAWQFFLLWSPDFR³PKMKASSLAFSLLSA⁴AFYLLWTPSTGLK⁵TNLGSCVIATNLQ⁶EIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK⁷PANRCCLLRHLLRLYLDRVFN⁸YQTPDHYTLRKIS
SLANSFLT⁹IKKDLRLSHAHMTCHCGTEAMKYSQILSHFEKLEPQAAVVKALGELDILQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCGCGATCCAGCC
 TAGCGTGTCCACGATGCGCGGCTGGGCTCCGGGACTTCGCTACCTGTTGGTAGCGATCGAGGTGC
 TAGGGATTCGGGTCTTCTCTCGGGGATTCCTCCGGCTCCGGTTCGTTCTCTGCCAGAGCGGAA
 CACGAGCGGAGGCCCGCAGCGCCGGAHCCCTCGGCTGGAGCCGATTCTAACTGGACACCGCTGCC
 ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTTGAGAGATGATTTTGTGTTTG
 GGTCAAAGGGTGTGAAATTTATGCCCTACACAACTTACCTTGTGGAAGAGGAGCATCTCACAGT
 TTTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTCGAATCAAGGCAATGATGACGGGAG
 CCTTCCTGGCTTGTGACGCTCATCAGGAACCTCAATTCTCCTGCACCTGCTGGAAGACAGTGTGA
 TAAGACAAGCAAAGCAGCTGGAAAAAGATAGTCTTTTATGGAGATGAACCTGGGTTAAATTA
 TTCCCAAAGCATTGTTGGAATATGATGGAACAACCTCAATTTTCGTGTCAGATTACACAGAGGT
 GGATAATAATGTCAGAGGCATTGGATAAAGTATTAAGAGAGGAGATTGGGACATATTATATCC
 TCCACTACCTGGGCTGGACACATTTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGGCGAG
 AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
 GACGCCCTTACCCAATTTGCTGGTCTTTTGTGGTGACCATGGCATGCTGAAACAGGAAGTCACG
 GGGCCTCTCCACCGAGGAGGTGAATACACCTCTGATTTTAACTCAGTTCTGCGTTTGAAGGAAA
 CCGGCTGATATCCGACATCCAAAGCAGCTCCAAAGACGGATGGCTGCGACACTGGCGATAGC
 ACTTTGGCTTACCATTCCAAAGACAGTGTAGGGAGCCTCTTATTCACAGTTGTGGAAGGAAGAC
 CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG
 AATGTGCCCTCATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTCAGAAAGATTGTCATGG
 GAACCTGGATCAGACTGTACTTGGAGGAAAGACATTCAAGAGTGTCTTATCAACCTGGGCTCCAAAG
 TTCTCAGGCGAGTACTGGATGCTCTGAAGACGCTGAGTGTCTGCTGAGTGCAAGTGGGCCAGG
 TTCTCACCCTGCTCTGCTGACGCTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAGATGCCA
 GTCTCATCTCTGGGTTTCTCTGCTCTTTATTGGTGATCTGTTCTTCTGGCGCTTCAAGT
 CAITGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCTCTCTGGCTGGCGGACAGCT
 GCCTTCTGTTTACCAGACTCTGTTGAACACTGTGTTGTGTCAGTGTGGCAGTGTGGCAGTGCCCTGGAC
 AGGGGGCTCAGGGAAGGACGTGGAGAGCCTTATCCAGGCTCTGGGTGTCGCCAGACAGGTG
 TTCACATCTGTGCTGTCTGAGTGCAGTGCCTCAGTTCTTGGAAAGCTAGTTCCTGCGACTGTTAC
 CAAGGTGATTGTAAGACCTGGCGGTCACAGAGAGAACAGGCCCCCAGCTGAGGGGGTGTGTGAA
 TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGAGAAAGAGACATCGGCCTGSA
 CACTCAGGAGGCTCAAAGAGAGACTTGGTGCACACCTATCTGCCACCCCGAATGCATCCT
 GCCTCATCAGGTCAGATTCTTTCACAGGCGGACGTTTCTGTGGAATCTTAGTCTTGGCC
 TCGGACACCTTCATCTGTTAGCTGGGAGTGGTGGTGAGGCACTGAAGAGAGAGCGGATGGTCAC
 ACTCAGATCCACAGAGCCAGGATCAAGGGACCACTGCAGTGGCAGCAGGACTGTTGGGCGCCC
 ACCCAACACTCGCACAGCCCTCATCCCTCTGGCTTGAAGCCTCAGAGGCCCTGTGCTGAGTGT
 CTGACCGAGACACTCAGAGCTTGTCTATCAGGACACAGGCTTCTCGGAGGCGAGGATGATCTGTG
 CCAGCTGTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAAATGATACCTAG
 CTGCACACAGTATGTAGTTACCAAAAGATAACGGCAATTAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPFAPEPSAGASSNWTLLPPLF
SKVYIVLIDALRDDFVFGSGKGVKMPYTTYLVKGSASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSEVIRQAKAAGRIVFYGDETWVKLFPKHFEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGWDILILHYLGLDHHIGHISGPNPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLGCDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCAGAGGCAAGCCTTCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACCTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACTACATACGCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCGGATGGTTCACTGCACTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTGAGAACAGGAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCAGTGAGGTCAGCGATTACGAAACTGCCCATTTGAACGCCCTTCTCGCTA
ATTGTGAACATAATTGTATAAAAAACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPNIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVDSGNLIAPDKNYIRPEI
FFALASSLSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGTGCCCTCCCTTTAA
TCCAGAGTCCCTGTCTCTCTGTCTGTAGGAGTGCCCTGTTGCCAGTGTGGGTGAGACAAGTTTG
TCCACAGGGGCTGTCTGAGCAGATAAGATTAAAGGGCTGGGTCTGTGCTCAATTAACCTCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCTACAGTCAGCACCATGCTGGGCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTTCCACGAGCAAAGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGTCTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAAGTCTAGTGTGGGAATTTGAAGACGACATTGACAAGTGCATTTCGAAGAAAGCACAG
AGCTGAACAATACTTTACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAGACCTGCTTGGAGGGATTCCACTGAGTGAACCCACTCACAGGCTTGTCCATGT
GCTGTCCACATTCGTTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTGTGTTATCTCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAAATCC
ACACATCTTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTTTAAGAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGSLWALLLLLSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAAGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGSCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAAGATCAAGTGAACCGGAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGGCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCAACCGCTGCATCAATGCCA
CCCAGGCGGGAACCAAGGGGGAGTCCAGAAAGCCAGACAAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCCAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCAACATGCACCAGCCAGTGCTCCTCTGCCCTTCGGCTTTGATCTGGCTCATGGTGAAT
AAGCTTGCCAGGAGGTGCGAGTACAGAGCGCAGAGCGAGCAAACTCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCTGCGCTTCIGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAATGCTTGTAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCAAACACTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGCTTGCTTAACCCAGGTTTTCTGTCTC
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTGAATAAATGAAACACCTGA

FIGURE 146

MRKHLSSWWNLATVCMLLFSHL SAVQTRG I K H R I K W N R K A L P S T A Q I T E A Q V A E N R P G A F I K Q G R K
L D I D F G A E G N R Y E A N Y W Q F P D G I H Y N G C S E A N V T K E A F V T G C I N A T Q A A N Q G E F Q K P D N K L H Q Q
V L W R L V Q E I C S L K H C E F W L E R G A G L R V T M H Q P V L L C L L A L I W L M V K

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

[illegible][illegible]

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKPLSDSGGTCE
GLFVEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 149

GTCTCOGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCOCTCTACCTGGAGACTTGAC
 TCCCGGCGGCCCCAAACCTGCTTATCCCTTGACCGTCGAGTGTGACAGATCCTGCACCGGCCAGTCC
 CGGCCCTCTCCCGCCCAACCCACCTCTCTGCTCTCTCTGCTTTTACTCTCTCTTTTATTATCATA
 ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCCAAGCCGAGCGTGAAGAATGGGGTT
 CCTCGGAGCGGCACCTTGGATTCTGGTGTAGTGTCTCCGATTCAAGCTTTCCCAACCTGGAGGAA
 GCGAAGACAAATCTCTACATAATAGAGAATTAAGTGCGAAGAACCTTTGAATGAACAGATTGCTGAA
 GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT
 TGTGTGAATCTGAACTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAAGAAAGACAATCTA
 TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGAAGATGTTGATTCAACCAAGAAATCGAAAACTG
 ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
 TCAACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
 AAAATGACAGAGCGGTGTTTGACAAAGATTGTTCTAAACTACTTAATCTCGGCTTTATCACAGAAAGC
 CAAGCACATACACTGGAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCACAA
 TTATGAGGAGGATCCCAATTAAGCCCAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAG
 TGACTCCAAATGGCAGCAATTCAAGTGGTCTTGCTAAGGGAGAAAACGATGAACAGTATCTAACACA
 TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAATCTCA
 ATATTTCCCAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAGCAAAAGAGAAAGAAA
 CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCT
 CCAGAAAGAGGTGTTTCTACCTTGAAAACCTTGGATGAATGATTGCTCTTCAGACCAAAAACAAGCT
 AGAAAAAATGCTACTGACAAATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
 CAGACAGTACCAGGAAGAAGCAGCTAAGATGGAAGAAGAAATTTGAAGCTTGAAGGATTCCACAAA
 GATGATACTCCAACCCAGGAGGAAGACAGATGAACCCAAAGGAAAAACAAGAGCCTATTGGAAGC
 CATCAGAAAAAATATTGAATGGTTGAAGAACATGACAAAAAGGAAAAATAAGAAGATTATGACCTTT
 CAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGATCCTTGACAAAGAA
 GAGCCGAGGCCATCAAGCGCATTTATAGCAGCTGTAAAAATGGCAAAAGATCAGAGATCTTTCAA
 CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATCTCTGTGATTAAATTTTTTGACCCAAAG
 GTTATTAGAAAGTGTGAAATTCAGTAGTTAACTTTTACAGTGGTTAAAJACATAGCTTCTTCCC
 GTAAAAACTATCTGAAAGTAAAGTTGTATGTAGCTGAAAAAAGAAAAAAGAAAAA

FIGURE 150

MGFLGTGTWILVLVLPQAFKPGGSQDKSLHNRELSAERPLNEQIAEAEEDIKKKTYPPENKPG
QSNYSFVDNLNLKKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIYSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIEKVTMAAIQDGLAGENDETVSNTLTLTNGLE
RRTKTYSEDNFEEQLQYFPNIFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISP EEGV
SYLENLDEMIALQTKNLEKNATDNISKLPAPSEKSHEETDSTKEEAAKMEKEYGSLKDDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGNKEDYDLKMRDFINKQADAYVEKGILLK
EEAEAIAKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCTGTGGAGCTCAA
 GATGCTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGTGTCTTATCTCG
 ATATAAACCAAGCTTTAGCTGGAGGGCTGCATGCAGGGAAGTTCATTAAAGGTGAAGAGATCAGC
 GTGGTCCCAATCGGTGGCTGGATGCCAGCTGTCCCCCGTCATCCTGGGTGCCAGGTGGGAAG
 CCAAGTGCCTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
 AGCTCTATCTTGGTGCCAAAGGAATCCAAGAGCTTCAACCTTCTACCGCGGGGACATGGGGCTCACC
 TCCAGCTTCGAGTCGGCTGCCTACCCGGCTGGTTCTGTGCACGGTGCCTGAAGCGATCAGCC
 TGTCAAGACTCACCAGCTTCCCAGGAATGGTGGCTGGAATGCCCCATCAGAGACTTCTACTTCC
 AGCAGTGTGACTAGGGCAAGCTGCCGCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
 GAGTGGAGGAGACCCATGGCGACAATCACTCTCTGTCTCAGGACCCCCACGCTCTGACTTAG
 TGGGCACCTGACCACCTTTGTCTTTCTGGTCCAGITTTGGATAAATTCTGAGATTGGAGCTCAGT
 CCACGGTCTCCCCACTGGATGGTGTCTGCTGTGGAACCTTGTAAAACCATGTGGGGTAAA
 CTGGGAATAACATGAAGAAGATTTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATCTGTCT
 TAATGGTAAGTGAAGAAGTGTACCCTGAGGCCCGCAGGCCAACCCATCCCCAGTTGAGCCCTTATA
 GGGTCAGTAGCTCTCCACATGAAGTCCCTGTCACTCAACACTGTGCAGGAGGGGAGGTGGTCTATA
 GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCTATA
 TGTACCTTTCTATCTCTTCCCTCATCATCTGTGTGGGCAATGAGGAGGTGGTGTATGCAGAA
 GAAATGGCTCGAGCTCAGAAGATAAAGATAGTAGGGTATGCTGATCCCTTTTAAAAACCCAA
 AGTACAATCAAAATCCAGAGTCTGGTCTCTATTCCCATGAAGGAGTGTCTGATGACATATTGAGA
 AGACCTACTTACAAGTGGCATATATTGCAATTTATTTTAAATGAAGATACCTATTATTATAT
 TCTTTATAGAAAAAGTCTGGAGAGTTTACTCTCAITTTGAGCAATGTGAGGGTGGGAGTAT
 AGGTGATTTTCTTTTAAATCTGTTAAATTTATCTGATTTTCTTCAATTTTCTCAATGAAGATGA
 ATCTCTGTATAAAAATAGAAAAGAAATTAATCTTGAGGTGAAGCAGAGCATCATCTCTGA
 TTGTCCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATGAATGAGCTCTGCTGCTCTGTGGT
 TTGTAGTAGTGATCAGGAACAGATCTCAGCAAGCCACTGAGGAGGAGGCTGTCTGAGTTTGT
 GTGGCTGGAATCTCTGGGTGAAGAACTTAAAGACAAAAATCATCTGGTAAATCTTCTCTAGAG
 GATCACAGCCCTTGGGATTCCAAGGCATTTGATCCAGCTCTCAAGAGGCTGCTGTACTGTTGA
 ATTGTGTCCCCCTCAAATTCACATCTTCTTGAATCTCAGTCTGTGAGTTATTGGAGATAG
 GTCTCTCAGATGTAGTTAGTTAAGACAAGGTCTATGCTGGATGAGGTAGACCTAAATCAATAT
 GACTGGTTTCTTGTATGAAGAGGAGGACACAGAGACAGAGGAGAGCCGGGAGAGCTATATA
 AAGATGAAGGCAGAGATCGGAGTTTGACGCCACAGCTAAGAAACACCAAGGATTTGTGCAAC
 ATCAGAAGCTTGGAGAGGCAAGAGAAATTTCTCCCTAGAGGCTTTAGAGGGATTAAGCGCTCTG
 CTCGAACCTTAATCTCAGACTTCCAGCTCTCTGAAGGAGAAAGATAAATTTCCGCTGTGTTTAA
 GCCACCAAGGATAATTTGGTTACAGCAGCTCTAGGAAACTAATACAGCTCTCAAAAATGATCCCTGT
 CCTCTCGGTGTTTACATTTCTGTGTGTGCTCCCTCCCAATGTCAAGAGTTGTCTTTGTGAGCAA
 TGAATATGSCAGAGTGATGGCAATGCCACTTCCAAGATATAGGTTATAAAGACACTGCAGCTTC
 TACTTGAAGCCCTCTCTCTCTGCCACCCAGCCGCCCAATCTATCTTGTGCTCACTGCTCTGGGG
 AGAGTCACTGCCATGCTATGACGAGGCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCTCT
 GGGTCAAGGCCATATAGTGAAGCTAAGAGCAGAGACTCTGTGAGATATCTGATGTTTGTGTTT
 AGTGTCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATCAGAGAAAGAG

60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 152

MVLGGALCFRMRKDSALKVLYLHNNQLLAGGLHAGKVIKGEISVVPNRWLDASLSPVILGVQGGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSPTTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGWGNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGCTGCAATGCGCCG
 CCTGCAGAAATCTGTGAGCTCTTTCCCTTATGGGGACCCCTGGCCACCAAGCTGCCTCCTCTCTTGG
 CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAAC
 TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
 CAACACAGACGTTGCTCTCATTTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
 ATCTGATGAAGCAGGTGCTGAACCTTCACCCCTTGAAAGAGTGCTGTTCCCTCAATCTGATAGGTTT
 CAGCCTTATATGCAGGAGGTGGTGCCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
 TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGA AAAAGC
 TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGACTGGATTGTCTGTTTATGCTCTGAGAAAT
 GCCTGCATTGCAACAGAGCAAAAGCTGAAAAATGAATAACTAACCCCTTTTCCCTGCTAGAAATAA
 CAATTAGATGCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG
 ATGGGTGGATTCCAATGAACCCCTGCGTTAGTTACAAAGGAACCAATGCCACTTTTGTTTATA
 AGACCAGAAGGTAGACTTCTAAGCATAGATATTTATGATAACATTTCAATTGTAAGTGGTGTTC
 TATACACAGAAAAAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
 TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
 TGTATTTATTTATTTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTTATTTAT
 AGAAACATCATTCGATATTGCTACTTGAGTGAAGGCTAATATTGATATTTATGACATAATTAT
 AGAGCTATAACATGTTTATTGACCTCAATTAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAAATAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGTTCAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCGAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCGCTTTTCCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCTAGAGGCTGCTAG
GCCAACCCGCCACCCAGAGTCTCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCAGGCCGT
TGCCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGGCACTCGGA
GCTGCTCTACCAACAACAGACTGTCTTCTACAGCGCGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGCGGCCCGCT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCCTGTGAAGTGTCTGTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAACCTGCATTTCTGCACATTTTGAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGCTCTGTCAATTTCTCTCAGGAAAGGTTTCAAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCGTACCCCTG
GCCCAGCACAGGCATTTCTAGATATTTCCCCCTTGCTGGAGAAGAAGAGCCCTGGTTTTATT
TGTGTTGTTACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTAAAAATGAAAA

FIGURE 156

MRERPRIDGESSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGGDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPNRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVPRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

COGGCGATGTCGCTCGTGTCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCGIACCCCGAGAGCC
GACCGTTCATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACATAGTGTGCAACAGGGGACTATTCA
ATTTGATGAATGTAGCTGGGTACTCCGGCAGATGCCAGCATCCGCTTGTGAGGGCCACCAA
GATTTGTGTACGGGCAAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTCATTGGGGCCATAATATTCCTAATGCAAAATATGAATGAAGATGG
CCCTTCCATGTCGTGAATTCACCTCACCAGGCTGCCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGGAAAGCCTGTGGGATCCGAACATCACTGCTTGTAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACACCACCTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCGAGTGACTCCATATTTTCCCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCAACACAGGCGTCCCTTTCCTCT
GGATAACAACAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTGCTGGTGGCCA
CATGGTGCTGTGGCAGGGATCTATCTAATGTGGAGGCAGGAAGGATCAAGAACTTCCCTT
TCTACCACCACACTACTGCCCCCTTAAAGGTTCTTGTGGTTTACCCTATCTGAATATGTTCCA
TCACACAATTTGTACTTCACTGAATTTCTTCAAACCATTCGAGAAGTGAGGTCACTCTGAAA
AGTGGCAAAAAAGAAAAAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCTCAAAAGAAAGCA
GCAGACAAAGTCGTCTTCTTCTTCCAAATGAGCTCAACAGTGTGTGCGATGTGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTTCCCTTGGCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATGTATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCATGCTTT
CTGTGCAGAACCTTCCATGTCAAGCAGCAGGTGTCAACAGGAAAAAGATCACAAGCCTGCCACG
ATGCGTGTGCTCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSEFTGPSPEWMLQHDLPGLDLADLVEFVTTTSVATGDYSILANVSVW
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRFSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNE DGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNF TTTPLGNRYMALIQH
STIIIGFSQVFEPHQKKQTRASVVIPTGDSGATVQLTPYPPTCGSDCIRHKGTVVLCPQTGVPFPLDNK
SKPGGWLFLLLLLLVATWVLVAGIYLMWRHERIKKTSFSTITLLFPKVLVVPSEICFHHTICYFTEFL
QNHCRSEVILEKNQKKKIAEMGPGVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGPSSENSQDLFELA
FNLPCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCTGCATGGCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCGGTCCCATCCAGCAA
GAGACCCCTGGTCGTCCGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTCAGTAAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16a

MTVKTLHGPMVKYLLLSILGLAFLSEMAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTWTWDPNRYPPSEVVQAQCRNLGCINAQKGEDISMNSVPIQGETLVV
RRKHQGCVSFSQLEKVLVTVGCTCVTFVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAGTGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAACTACCCAGCAGCAGCCCCCTCCGCCCTCTCGAGAGGCTGAAGAGGGATTTC
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCCCTTGGGGGGGGCAGCAC
AGGGCCCTCAGGCCCTGGGTGCCACTTGGCACTAGAAATGTCCTGTGCCTTCTTGTCTGCTCT
TGGCACTGGGCCGAAGCCAGTGGTCTTCTCTGGAGAGGCTTGTGGGGCTCAGGACCTTACC
CACTGCTCTCCGGGCTCTCTCTGCCGCTCTGGGACAGTGACATACTTGTCTGCTCGGGACAT
CGTGCTGCTCCGGGCCCGGTGCTGGCGCTACGCACCTGCAGACAGAGCTGCTGCTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTCTGCTGCTGGCTGTCCACTTGGCGTGCATGGGCATGG
GAGAGGCTGAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCTTAGGAA
TGCCTCTCTCAGGCCCAAGTGTGCTCTCTCTCCAGGCTACCCCTACTGCCCGTGCCTCTGCTG
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGAGTGTGGTCACTGTGGGCTCTGTGATATGAC
TGTCTGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTACTCAGCCAGGTACGAGAA
GGAACTCAACACACACAGCAGCTGCTGCTGCTGCTGCTCAACGTGTGAGCAGATGGTGACA
ACGTGATCTGGTTCTGAATGCTCTCGAGGAGCAGCACTTGGGCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCGGTGGGCAAAAACCTGACTGGACCGCAGATCATTACCTTGA
CCACACAGACTTGGTTCCCTGCTCTGTATTCAAGTGTGGCTCTGGAACCTGACTCCGTTAGA
CGAATCTGCCCCCTCAGGGAGGAGCCCGCGCACACACAGAACTCTGGCAAGCCCGCGACTG
CGACTGCTGACCTGCAGAGCTGGCTGCTGAGCACCCTGCTGCTGCCGCGAGCGAGCGCACT
GTGCTGGCGGCTCCGGGTGGGGACCCCTGCCAGCACTGGTCCACAGCTTCTCTGGGAGAAAG
TCACTGTGGACAGGTTCTCGAGTTCCCATGCTGAAAGGCCACCTTACTCTGTGTTTCAAGGTG
AAGAGTCCGAGAAGCTGCAGCTGCAGAGTGTCTTGGGCTGACTCCCTGGGGCTCTCAAGAA
CGATGTGCTACTGTGGAGACACAGGGCCCCAGGACAACAGATCCCTCTGTGCTTGGAAACCA
GTGGCTGTACTTCACTACCCAGCAAGCCCTCCAGAGGCGAGCTGCTTGGAGAGTACTTACTA
CAGAGCTGCACTCAGGCGAGTGTCTGCACTATGGAGCATGACTTGGAGCGCTATGGGCTG
CCCCATGGCAAAATACATCCACAAGCGCTGGGCCCTGTGTGGCTGGCTGCTTACTCTTGGCG
CTGCGCTTCCCTCATCTCTCTCTCAAAAAGGATCAGCGAAAGGGTGGCTGAGGCTCTTGA
CAGAGCTCCGCTCGGGGGCGCGCGCAGGGCGCGCGGCTCTGCTCTCTACTCAGCGATGA
CTCGGGTTTCAGGCGCTGCTGGGCGCTGCTGCTGCGGCTGTGCGAGCTGCCCTGCCGCTGG
CGGTAGAGCTGTGAGCGCTGTGAGCTGAGCGCGCAGGGGCGCGGCTTGGTTTCAAGCGCAG
CGCGCGCAGACTGCAGAGGCGCGCTGCTGGTCTTGTCTCTCTCCCGGCTGGCTGGGCT
GTGCGAGGAGTGGCTACAGGATGGGTGTCCGGGCGCGGGCGCAGCGCGCAGCGCTTCC
GCGCTCGCTCAGTGTGCTGCTGCCGACTCTTGCAGGGCGCGGGCGCGCGCAGCTTACGTGGG
GCTGCTTGCAGAGCTGCTCACCAGGAGCGGCTACCGGCTTTTCCGACCGTGGCGCTCTT
CACACTGCGCTCCCACTGCGAGACTTCTGGGGGCTTGCAGCAGCTCGCGCGCGGCTTCC
GGCGGCTCCAAGAGAGCGGAGCAAGTGTCCGGGCTTTCAGCGAGCTTGGATAGTACTTCC
CATCCCCGGGACTTCCGCGCCGGAGCGGGGTGGGACAGGGGCGGAGCTGGGGCGGGGA
CGGACTTAAATAAAGGCAGAGCTGTTTTCTAAAAAA

FIGURE 162

MFVWFLLSLALGRSPVVLSERLVGPQDATHCSFGLSCLWSDSILCLPGDIVPAGPVLAPTHQTELV
 LRCQKETDCDCLLRVAHLAVHGHWEPEDEERKFGGAADSGVEEPNALSQAQVVLSPQAYFTARCVLLEV
 QVPAALVQPGQSVGSVYDCFEAALGSEVRINISYQPRYEKELNHTQQLPALPWLINVSADGNVHLVLNVS
 EEQHFGLSLYWNQVQGPQKPRWHKMLTGPILITLNNHTDLVPCLCIQVWPLEPDSVRTNICFPREDPRAHQN
 LMQAARIALLLTQSWLLDAPCSLPAEAALCWAPGGDPCQPLVPPLSWENVTVDRKLEFPLKGFHPLNCVQ
 VNSSERLQLQECLWADSLGLKDDVLLLETRGPDNRSLCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
 GQCQLWDDDLGALWACPMOKYIHKRWALVWLACLLFAAALSLLLLKKDHAGWLRLLKQDVRSGAAARG
 RAALLYSADDSGFERLVGALASALCQLFLAVAVDLWSRRELSAQGPVAFWFAQRERQTLQEGGVVLLFSP
 GAVALCSRWLQDGVSGPGAHGPHDAFRASLSCVLEDFLQGRAPGSYVGACFDRLHHPDAVPALFRTVPVFT
 LPSQLPDLFGALQPPRAPRSGRLQERAEQVSRAQLQPALDSYFHPPTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
 692-697, 696-701, 700-705

FIGURE 164

MRTLTLTILTVGSLAAHAPEDPSDLLQHVKPQSSNFENILTWSGPEGTPDTVYSIEYKTYGERDW
VAKKGQQRITRKSCNLTIVETGNLTLYYARVTAVSAGGRSATKMTRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCVRKTLTPDRTWYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTQPLRFIQEHVLIIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVFPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFFPYAPQAISKVPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRITSDPNVLHSGEETPQYLGKQLPLLSSVQIEGHMSLPLQPPSGPC
SPSDQGPSPWGLLESVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAACATG
CTGCGGGCGCGGGGCTGCTCTTCTGGCTGTTCTGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTTGGGTGGA
AGTGTTGAACACAGTTTTGGATATTTTCCAAAAGATTGTATCAAGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGSATCTTTGGAACCTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCTGTGCCTGAGCCCGAGGCATTCAGAGCTGATTGAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCTGTGGACACTTTTGAAGAAATTCGACAGATAA
ATTGAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGC GTTATT
CATTACAGCAAGGATTTCGTTGGCATCAAATCTAAGTTTGTGTTTACAAAGATTGTTTTTAGTA
CTAAGTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTTCCCTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLEFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEECMSMLMYRGKALEDFTGPDRCRFVN
FKKGDDVVVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELDSVPESKKAEEVSQHREKSPESRGRELDPVPEPEAFRADSEDGEGA
FSESTEGLGQGPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQSGGQCVIHYSGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGACAGGCGCGGGAAGAGAAGCAAGCGGC
AACCGTGTGGTCCAAAGCCGGGGCTTCTGCTTGCCTCTAGGACATACACGGGACCCCTTACTTC
AGTCCCCCAAACCGCGCACCTCTGAAGTCTTGAACCTCAGCCCGCGACATCCACGGCGGCGCACGG
CGCGCGACGGCGCAGGTCCCGCGCGAAGCGATGCGCGCAGGGGTCGGGCACTGGGCTCTGGGC
GGCGGGAGTAGGGCCCCGCGAGGAGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCGCATGAGCGCGGTGG
TCTGCTGCTGCTGGGCGCGCGCTGCTGCTGCGGCGCAGGAGCCTTCTGCCGCGCGTGGTGCAG
GGCCAAAGGTGTGTTTGTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCTGGCTTGTGAGAGTGAGGGAGGAGTCTCTCTCA
GCCTTGAAGATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGAAAACCTGACAAAACCCGGG
ACAGGATTTCTGATGGTGATTCTCGATAGGGCTTTGGAGGAATGGAGATGGGCAACATCTGG
TGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATCCCGAGTACCGAAACTGGTACACAG
ATGAACCTCTCTGCGGAAGTGAAAAGTGTGTTGTGATGATACCAACCAACTGCGCAATCCTGGC
CTTGGGGGCTCCCTACCTTTACCAAGTGAATGATGACAGGTGTAACATGAAGCACAAATATATTTG
CAAGTATGAACCAGAGATTAAATCCAAACAGCCCTGTAGAAAAGCCTTATCTTACAAATCAACAG
GAGACACCCATCAGAAATGTGGTTGTACTGAAGCAGGTATAATTCCAAATCTAAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAGGAAGAACAAAACTAGTCCAAACAGCTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAAGAAAGTGGCATGGAAGTATAAATAACTCATTGACTTGGTTCAGAAATTTGTAAATCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGAATGGCTTGAATCACAAGGATCTGC
AAGATGAAGTGAAGCTCCCTCTGAGGCAATATTAAGATAATTTTATATGTCTATTATTTCA
TTTAAAGAAATAGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGTATAAGGATGCACCCAA
ACTTCAAACCTCAAGCAATGAATGGCAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTAGAAGCAATCCTTTATTTCTTTCACTTTTCAATAGTTGTTATCTAGTCAATGTAA
TGATATTTGATTGAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAAA
ATGAAGTGTCTAATAATTTATTTTATGGCACTCTATTTTCAATACATGCTCTTTTGATTAAAG
AACTTTATTACTGTGTCACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATATTTCACTTTTCACTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTCAGA
AATAAGAGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAACATTTTACTTAGNGGCAAGAT
TGCTCAATTTCAATTTGTGCAAGCATGTGCCCTTATAATATTTTTAGCTTAAATTAACAGATT
TTGTAATTAATGTAACCTTTGTTAATAGTGCATAAACACTAATGCAAGTCAATTTGAACAAAGAG
TGACATACACAATATAAATCATATGTCTTCACAGGTGCGTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTGCCCACTAAACAAAGATGGTTGTTGGGGTTTGGG
ATTGACCTGGAGGCAGATAGTTGCAAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCAATGTGGTTGAGACAGGTGAATAGTCACTATCAGTGTGGAG
ACAGCACGACGACACAGACATTTAGGAAGGAAAGGAACACGAAATCGTGTGAATGGGTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGTTTGTCTGAGATAGAAAATGGTGGCTCTCTT
CTGCTCTTCTCTCCTAGTTTCTTCAATGCTTACGCTTCTTCTCTCAAGAGAAAGTTGTAACCT
CTGGCTCTCATATGTCCCTGTGCTCTCTTTAACCAATTAAGAGTCTTGTGTTCTCGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVSVSGQKVCFADFKHPCYKMAFYHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDWFGLWRNGDGTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNHYICKYEFEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKS KGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217